the
university of
connecticut
libraries

hbl, stx
QA 247.R52
Clifford numbers and spinors.
3 9153 00624028 9

QA/247/R52

# UNIVERSITY OF MARYLAND

Lecture Series No. 38

# CLIFFORD NUMBERS AND SPINORS
## (Chapters I - IV)

given by

Professor Marcel Riesz

Visiting Research Professor
University of Maryland
Professor Emeritus, University of Lund

Lectures delivered October 1957-January 1958

THE INSTITUTE FOR FLUID DYNAMICS
*and*
APPLIED MATHEMATICS

WILBUR L. CROSS LIBRARY
UNIVERSITY OF CONNECTICUT

QA
247
R52

Lecture Series No. 38

# CLIFFORD NUMBERS AND SPINORS
(Chapters I - IV)

given by

Professor Marcel Riesz

Visiting Research Professor
University of Maryland
Professor Emeritus, University of Lund

Lectures delivered October 1957-January 1958

512.815
R446c

QA
247
R52

# PREFACE

The presentation of the theory of Clifford numbers and spinors given here has grown out of thirteen lectures delivered at the Institute for Fluid Dynamics and Applied Mathematics of the University of Maryland. The work is divided into six chapters which, for the convenience of those readers who are only interested in certain parts of the material treated, are largely independent of each other. This arrangement has, of course, certain disadvantages such as repetitions and overlappings.

Only Clifford algebras over either the real or the complex field are considered.

Chapter I is mainly algebraic. In the first half we give the main features of Clifford algebras associated with an arbitrary metric, regular or singular. The connection between Clifford algebras and tensor algebras, in particular between Clifford numbers and skew symmetric tensors, is treated very thoroughly. The second half of this chapter, beginning with section 1.12 is considerably more difficult than the first one. However, it is not needed for the understanding of later chapters. Sections 1.12-1.17 carry out some beautiful ideas of Chevalley in a slightly modified way. We treat derivations in graded algebras and obtain a representation of a Clifford algebra as the quotient algebra of a tensor algebra modulo a certain ideal. In sections 1.18-1.19 we show how derivations lead to interesting results also in Clifford algebras, in spite of the fact that such an algebra is not graded.

Chapter II is entirely geometrical. It originates from Hamilton's idea to obtain a rotation by means of two reflections. This leads to the E. Cartan-Dieudonné theorem telling that every isometry can be obtained by means of successive reflections. The geometrical operations in question can in an efficient and elegant way be expressed in the language of Clifford algebra. We thus come in contact with the spinor representation of isometries.

The main object of Chapter III is the canonical decomposition of a given isometry in a space over the real field and with a metric of Euclidean or Lorentz signature. However, this chapter contains a general introduction about minimal polynomials of linear transformations, decomposition of a space into subspaces corresponding to relatively prime factors of the minimal polynomial, orthogonal complement, radical etc. Among the possible isometries of a space with Lorentz signature the often neglected or forgotten parabolic rotations are carefully treated. Clifford algebras are entirely missing in this chapter.

Chapter IV begins by the theory of infinitesimal isometries. These are those linear transformations which are skew-symmetric with respect to the metric in question. In the Lorentz case special attention is paid to the parabolic infinitesimal isometries. In section 4.11 the Clifford algebras reappear due to the connection between skew-symmetric transformation and Clifford bivectors. It is shown that in the case of a real field and a metric with Euclidean or Lorentz signature every isometry connected with the identity can be expressed by the exponentials of a single infinitesimal isometry or a single Clifford bivector. It is also shown that this property is not shared by any other signature. In section 4.16 we briefly consider the Lie algebras formed by Clifford bivectors or by bivectors and vectors. In section 4.17 we write Maxwell's equations in Clifford notations. We also interpret the Lorentz force (ponderomotive force) as a local infinitesimal isometry and show that in the case of a plane wave the four-velocity of a moving charge is subjected to infinitesimal parabolic rotations.

Chapter V treats some special features of four dimensional real Lorentz space, i.e., the space-time of special relativity. We deal largely with properties which neither belong to any Lorentz space of some other dimension, nor to a space of some other signature and of any dimension. The direct sum of the field of the real numbers and of the space of real pseudoscalars forms a field, isomorphic with the field of the complex numbers. Over this field we construct a geometry whose elements are pencils of bivectors. We also tell some interesting things about the Maxwell-Minkowski stress-energy tensor.

Chapter VI gives a short introduction into the theory of spinors. Applications of this theory to the Dirac equation are also given.

I should like to express my gratitude to my friend John Horvath for the invaluable help he has given me in the endeavor to work out these enlarged Lecture Notes. Chapter II and the Appendix to Chapter IV are edited by him. His mastery of modern theories has helped me to overcome many obstacles. Without his energy and his enthusiasm the task would have been abandoned at an early stage.

My thanks are also due to several colleagues, in particular to Dr. Martin Pearl, for the kindness they showed to me in revising and correcting my English.

## Chapter I

## Clifford Numbers

1.1 We are going to construct associative but not commutative algebras over the real or the complex field. A natural approach to Clifford numbers arises from the following problem to write the quadratic form $\sum_{i=1}^{n} x_i^2$ as the square of the linear expression $\sum_{i=1}^{n} x_i e_i$, that is

(1.1.1) $$\left(\sum_{i=1}^{n} x_i e_i\right)^2 = \sum_{i=1}^{n} x_i^2 .$$

The same problem arises in the relativistic quantum mechanics of the electron in a slightly different form. For the moment let us replace the wave operator by the closely related Laplace operator

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} ,$$

in the statement of the problem. We want to find a differential operator $\sum_{i=1}^{n} e_i \frac{\partial}{\partial x_i}$ of the first order such that

$$\left(\sum_{i=1}^{n} e_i \frac{\partial}{\partial x_i}\right)^2 = \Delta .$$

Calculating the left hand side of (1.1.1) formally and using the distributive rule, we obtain

$$\sum_{i=1}^{n} x_i^2 e_i^2 + \sum_{i<1} x_i x_j (e_i e_j + e_j e_i).$$

It follows that (1.1.1) will be satisfied if

(1.1.2) $$e_i^2 = 1, \quad e_i e_j + e_j e_i = 0 \text{ for } i \neq j,$$

or equivalently if

(1.1.3) $$e_i e_j + e_j e_i = 2\delta_{ij},$$

where $\delta_{ij}$ is the Kronecker symbol defined by $\delta_{ii} = 1$, $\delta_{ij} = 0$ if $i \neq j$.

Now let $x = \sum_{i=1}^{n} x_i e_i$ and $y = \sum_{i\ 1}^{n} y_i e_i$ be two vectors expressed in terms of a given basis $e_1, e_2, \ldots, e_n$ of orthogonal unit vectors in an n-dimensional Euclidean space. We write the Euclidean scalar square $(x \mid x)$ and the scalar product $(x \mid y)$

$$(x \mid x) = \sum_{i=1}^{n} x_i^2, \qquad (x \mid y) = \sum_{i=1}^{n} x_i y_i.$$

Identifying the basis vectors $e_i$ with the symbols $e_i$ in (1.1.1), we may write

(1.1.4) $$x^2 = (x \mid x).$$

Clearly this equation alone completely replaces the detailed relations (1.1.2) or (1.1.3).

The preceding problem can immediately be generalized. Consider an n-dimensional vector space $E$ with a general metric form (scalar product) defined by

(1.1.5) $$(x \mid y) = \sum_{i,j=1}^{n} g_{ij} x_i y_j,$$

where $g_{ij} = g_{ji}$. If $g_{ij}=\delta_{ij}$ we are in the Euclidean case considered above. If we put $g_{11} = 1$, $g_{ii} = -1$ for $i = 2,\ldots, n$ and $g_{ij} = 0$ for $i \neq j$, we obtain the so-called Lorentz metric, which will be of fundamental importance in these Lectures and which defines for $n = 4$ the space-time of special relativity. Another interesting, though highly degenerate case is when all $g_{ij}$ are zero. This case yields the Grassmann algebra or exterior algebra.

We again ask for a rule which yields (1.1.4), where the right hand side is to be interpreted according to (1.1.5) and obtain this time the relations

$$(1.1.6) \qquad e_i e_j + e_j e_i = 2g_{ij},$$

which obviously generalize (1.1.3).

From $x^2 = (x \mid x)$ we infer the polarized form of (1.1.4), namely

$$(1.1.7) \qquad (x \mid y) = \tfrac{1}{2}(xy + yx).$$

Indeed since (1.1.4) holds for any vector, we have

$$(x + y \mid x + y) = (x + y)^2,$$

or, using the relation $(x \mid y) = (y \mid x)$,

$$(x \mid x) + 2(x \mid y) + (y \mid y) = x^2 + xy + yx + y^2.$$

Subtracting and recalling that $(x \mid x) = x^2$, $(y \mid y) = y^2$, we obtain (1.1.7). Two vectors $x$ and $y$ are said to be orthogonal if $(x \mid y) = 0$. It follows from (1.1.7) that if $x$ and $y$ are orthogonal then $xy + yx = 0$. This last relation is expressed by saying that $x$ and $y$ anticommute.

We are thus led to an algebra C generated by the basis elements $e_1, e_2, \ldots, e_n$ of the vector space E. This algebra contains all the products of such basis elements and also all the linear combinations of such products. The product of two of these linear combinations is defined by distributivity. Since the square of every vector is a scalar, this algebra necessarily contains all the scalars, except in the Grassmann case, where they are included by definition. The scalars are assumed to commute with every element of C. The multiplication in C is not commutative, but as we show later it is associative. The algebra C is called the Clifford algebra associated with the metric (1.1.5). The elements of C are called Clifford numbers.

1.2. Until section 1.7 we consider a metric form (1.1.5) for which $g_{ii} \neq 0$ for $1 \leq i \leq n$ and $g_{ij} = 0$ for $i \neq j$. In this case every element $e_i$ possesses an inverse in C, namely the the element $e_i^{-1} = g_{ii}^{-1}\ e_i$. Furthermore two different $e_i$ are orthogonal to each other and thus anticommute, i.e.

$$e_i e_j + e_j e_i = 0, \qquad i \neq j. \tag{1.2.1}$$

It follows from (1.2.1) that if we want to obtain a basis of the algebra C, it is enough to consider among all the products $e_i e_j$ of two basis vectors only those for which $i < j$. More generally, if we have a product of the form $e_{\beta_1} e_{\beta_2} \cdots e_{\beta_p}$, where the $\beta_i$ are distinct $(1 \leq \beta_i \leq n)$, then we clearly have

$$e_{\beta_1} e_{\beta_2} \cdots e_{\beta_p} = (-1)^{\sigma}\, e_{\alpha_1} e_{\alpha_2} \cdots e_{\alpha_p},$$

where $1 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_p \leqslant n$ and $\sigma$ is the number of inversions in the permutation $(\beta_1, \beta_2, \ldots, \beta_p)$ of $(\alpha_1, \alpha_2, \ldots, \alpha_p)$. Thus for example $e_5 e_1 e_4 e_3 = - e_1 e_5 e_4 e_3 = e_1 e_4 e_5 e_3 = - e_1 e_4 e_3 e_5 = e_1 e_3 e_4 e_5$. Finally, if the same basis element $e_i$ occurs repeatedly in a product then, after ordering the factors according to increasing indices, $e_i$ will appear at most to the first power since $e_i^2$ is a scalar. It follows in particular that every product can be reduced to one containing $\leqslant n$ factors.

We have thus shown that the products

$$e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_p} \tag{1.2.2}$$

with $1 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_p \leqslant n$, $0 \leqslant p \leqslant n$, if proved linearly independent, form a basis for the algebra C. We state here once for all that among the products (1.2.2) the empty product, which stands for the element 1, is always included. We shall show the linear independence in section 1.4.

We shall usually denote a system $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ with $1 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_p \leqslant n$, $0 \leqslant p \leqslant n$ by a capital letter, say A, and the corresponding product (1.2.2) will be abbreviated by $e_A$. If $p = 0$, the product is empty and denotes the scalar 1.

**1.3.** It will be sufficient to prove the associativity of the multiplication in C for the Euclidean case. By what precedes, it is enough to show that $e_A(e_B\, e_C) = (e_A\, e_B)\, e_C$ for all elements of the system $\{e_A\}$ just constructed.

With a slight change of notation, let us write

$$E_{\rho_1,\rho_2,\dots,\rho_n} = e_1^{\rho_1} e_2^{\rho_2} \dots e_n^{\rho_n} ,$$

where the $\rho_i$ are taken modulo 2. With this notation, the multiplication of two such elements can obviously be written

$$E_{\rho_1,\dots,\rho_n} E_{\sigma_1,\dots,\sigma_n} = (-1)^{\sum\limits_{i>j} \rho_i \sigma_j} E_{\rho_1+\sigma_1,\dots,\rho_n+\sigma_n} .$$

Now $E_{\rho_1,\dots,\rho_n}(E_{\sigma_1,\dots,\sigma_n} E_{\tau_1,\dots,\tau_n})$ is easily seen to be

$$(-1)^{\sum\limits_{i>j} \rho_i \sigma_j + \sum\limits_{i>j} \rho_i \tau_j + \sum\limits_{i>j} \sigma_i \tau_j} E_{\rho_1+\sigma_1+\tau_1,\dots,\rho_n+\sigma_n+\tau_n}$$

from which the associativity clearly follows.

1.4. We shall prove in this section that the system $e_A$ constructed in 1.2 is linearly independent.

If $e_A = e_{\alpha_1} e_{\alpha_2} \dots e_{\alpha_p}$ is an element of this system, we have $e_j e_A = (-1)^q e_A e_j$ with $q=p$ or $q=p-1$, according as $j$ is different from all the $\alpha_i$ or $j=\alpha_i$ for some $i$. This shows that, when $p$ is even, $e_A$ anticommutes with every $e_j$ such that $e_j$ is one of the $e_{\alpha_i}$ and commutes with every $e_j$ which is different from all the $e_{\alpha_i}$, while, if $p$ is odd, the situation is the reversed.

It follows that *to every product* $e_A = e_{\alpha_1} e_{\alpha_2} \dots e_{\alpha_p}$ *with* $p \geqslant 1$ *there exists a basis vector* $e_j$ *which anticommutes with* $e_A$, except when the dimension $n$ of $E$ is odd and $e_A$ is the "pseudoscalar" $e_1 e_2 \dots e_n$.

We are now in the position to prove the independence of the elements $e_A$. Suppose first that $n$ is even. Consider a linear relation

$$(1.4.1) \qquad \sum_A c_A e_A = 0$$

among the elements $e_A$, where the coefficients $c_A$ are scalars. We want to show that all $c_A = 0$. Suppose there exists a $c_A \neq 0$, then multiplying (1.4.1) by $c_A^{-1} e_A^{-1} = c_A^{-1} e_{\alpha_p}^{-1} \dots e_{\alpha_1}^{-1}$, we obtain

$$(1.4.2) \qquad 1 + \sum_B c_B e_B = 0.$$

If (1.4.1) contained the single term $c_A e_A$, we have already the contradiction $1 = 0$. If not, we shall arrive at the same contradiction by a reduction process. Suppose that there is a $c_{B_o} \neq 0$. We know that there exists a vector $e_j$ which anticommutes with $e_{B_o}$. Multiplying (1.4.2) by $e_j$ from the left and by $e_j^{-1}$ from the right we obtain

$$(1.4.3) \qquad 1 + \sum_B c_B e_j e_B e_j^{-1} = 0$$

Adding (1.4.2) and (1.4.3) and dividing the sum by 2, we obtain the relation

$$(1.4.4) \qquad 1 + \sum_B \tfrac{1}{2} c_B (e_B + e_j e_B e_j^{-1}) = 0.$$

Since $e_j e_B e_j^{-1} = \pm e_B$, this expression is of the same form as (1.4.2), but contains at least one term less with nonvanishing coefficient, since $e_j e_{B_0} e_j^{-1} = -e_{B_0}$ and all the terms with zero-coefficient in (1.4.2)will also have zero coefficient in (1.4.4). By iterating this operation we arrive after a finite number of steps at the desired contradiction $1 = 0$.

If now n is odd, then, as shown above, there exists no $e_j$ which anticommutes with the pseudoscalar $e_1 e_2 \dots e_n$. Thus, in this case, the above construction leads either to $1 = 0$ or to

$$(1.4.5) \qquad 1 + \beta e_1 e_2 \dots e_n = 0.$$

We shall show that this last relation is also impossible. Consider the automorphism of C defined by $e_i \rightarrow -e_i$ for $1 \leqslant i \leqslant n$. Since $(-e_i)(-e_i) + (-e_j)(-e_j) = 2g_{ij}$, all relations in C hold also if we replace every $e_i$ by $-e_i$. In particular we have from (1.4.5)

$$1 + \beta(-e_1)(-e_2) \dots (-e_n) = 0,$$

and since n is odd

$$(1.4.6) \qquad 1 - \beta e_1 e_2 \dots e_n = 0.$$

Adding (1.4.5) and (1.4.6), we arrive at the contradiction $2 = 0$.

We have thus proved that the $2^n$ products

$$(1.4.7) \qquad e_A = e_{\alpha_1} e_{\alpha_2} \dots e_{\alpha_p}$$

with $1 \leqslant \alpha_1 < \alpha_2 < \dots < \alpha_p \leqslant n$, $0 \leqslant p \leqslant n$, form a basis of the algebra C, whose dimension over the field of scalars is hereby

$$\sum_{p=0}^{n} \binom{n}{p} = 2^n.$$

If for a fixed p we call $C_p$ the $\binom{n}{p}$-dimensional subspace of C spanned by the products (1.4.7) with exactly p factors, then C is the direct sum of the subspaces $C_p$. We identify $C_0$ with the field of scalars and $C_1$ with the vector space E. The subspace $C_n$ has dimension 1 and its elements are called pseudoscalars. If we set

$$C_+ = \sum_{p \text{ even}} C_p, \qquad C_- = \sum_{p \text{ odd}} C_p,$$

then C is the direct sum of the two subspaces $C_+$ and $C_-$, each of dimension $2^{n-1}$. Furthermore $C_+$ is a subalgebra of C and we have the relations

$$(1.4.8) \quad C_+C_+ \subset C_+, \; C_- C_- \subset C_+, \; C_+ C_- \subset C_-, \; C_- C_+ \subset C_-.$$

The elements of $C_+$ will be called even, those of $C_-$ will be called odd.

We introduce an important automorphism J of C called the main involution, lightly touched upon already above. For $u \in C_p$ we put $J(u) = (-1)^p u$ and extend J to C by linearity. By (1.4.8) we have $J(uv) = J(u)J(v)$, that is J is indeed an automorphism. Furthermore $J^2$ is the identity, which accounts for the name of J.

We also introduce an antiautomorphism $c \rightarrow \tilde{c}$ of C called the reversion or the main antiautomorphism. This is defined by writing $\tilde{e}_A = e_{\alpha_p} e_{\alpha_{p-1}} \ldots e_{\alpha_1}$ for the elements (1.4.7) and extending the definition by $c = \sum c_A e_A \rightarrow \tilde{c} = \sum c_A \tilde{e}_A$. Without reference to the basis, the reversion can be defined as a linear transformation of C into itself which leaves every scalar and every vector invariant and satifies $(cd)^{\sim} = \tilde{d}\tilde{c}$. Noting that $1 + 2 + \ldots + (p-1) = \frac{1}{2} p(p-1)$ we find $\tilde{u} = (-1)^{\frac{1}{2}p(p-1)} u$ if $u \in C_p$.

Written out in detail, this means that $\tilde{u} = u$, if $p \equiv 0,1 \pmod 4$ and $\tilde{u} = -u$, if $p \equiv 2,3 \pmod 4$.

1.5. We now turn our attention to the structure of the algebra C. We recall that the *center* of an algebra $\mathcal{A}$ is the subalgebra Z consisting of those elements of $\mathcal{A}$ which commute with all elements of $\mathcal{A}$.

Let us first suppose that *the dimension* n *of* E *is even*. We prove that *in this case the center* Z *of* C *reduces to the scalars* $C_0$. Indeed, the scalars commute with every element of C. Conversely, let $c = \sum c_A e_A$ belong to Z and suppose that there exists a non zero coefficient, say $C_B$, with non empty B. We know that there exists an $e_j$ which anticommutes with $e_B$. It follows that $c = e_j c e_j^{-1} = \sum c_A e_j e_A e_j^{-1} = \sum \pm c_A e_A$, where in particular the minus sign has to be taken in the term corresponding to A = B. Thus c would have two different expressions in terms of the basis $e_A$ in contradiction to the linear independence of the $e_A$.

*If* n *is odd*, then we can see in an entirely analogous manner that *the center of* C *is* $C_0 + C_n$, that is the subalgebra formed by all elements of the form $\alpha + \beta e_1 e_2 \dots e_n$.

Next let us recall that a two-sided *ideal* I of an algebra $\mathcal{A}$ with a unit element is a subset of $\mathcal{A}$ having the following properties:

(I) $a \in I$ and $b \in I$ imply $a \pm b \in I$,

(II) $a \in I$ and $c \in \mathcal{A}$ imply $ac \in I$ and $ca \in I$.

These two properties can also be stated in the following equivalent form:

(I') $I \pm I \subset I$,

(II') $\mathcal{A} I \subset I$, $I \mathcal{A} \subset I$.

Obviously an ideal is also a subalgebra of $\mathcal{A}$. The subset of $\mathcal{A}$ composed of the single element 0 is a two-sided ideal; it is denoted by (0) and is called the zero ideal. The whole algebra $\mathcal{A}$ also forms a two-sided ideal. Since this ideal is generated by the unit element, it is denoted by (1) and is called the unit ideal. Any ideal different from (0) and (1) is called a proper ideal.

Let us now return to the Clifford algebras C and let us first assume that $n = 2\nu$ is even. Then the only two-sided ideals in C are (0) and (1). This is expressed by saying that the algebra C is simple. To prove this, let I be a two-sided ideal in C, different from (0). Then I contains a non zero element $\sum_A c_A e_A$. Multiplying by a $c_B^{-1} e_B^{-1}$ such that $C_B \neq 0$, we obtain from (II) an element of the form $1 + \sum_A c_A e_A$ contained in I. Now the same reduction can be undertaken which yielded in 1.4 the proof of the independence of the basis $e_A$. By (I) and (II) every step again yields an element belonging to I. Hence we arrive at the conclusion that $1 \in I$, which means that $I = (1) = C$.

The fact that C is simple, has an important consequence, which we only state for the case of complex scalars. It is a well known general fact that a simple algebra $\mathcal{A}$ over an algebraically closed field is isomorphic to a total matric algebra $\mathfrak{M}$ over the field, that is to an algebra consisting of all square matrices of a certain order $m$ whose elements lie in the field. This implies that the dimension of $\mathfrak{A}$ is equal to the square of an integer, namely to $m^2$. Applying all this to the present case, C is isomorphic to the algebra of all $m = 2^\nu$ rowed square matrices whose elements are complex numbers; we give a direct proof of this in chapter VI. In particular if $n = 4$, $\nu = 2$, then $m = 2^\nu = 4$, that is $m = n$, just a coincidence.

Next we pass to the case where the dimension $n$ of $E = C_1$ is odd and consider first only complex scalars. Let us introduce the pseudoscalar

$$P = \varepsilon e_1 e_2 \dots e_n,$$

where $\varepsilon$ is determined in such a way that

$$P^2 = 1. \tag{1.5.1}$$

Explicitly we have

$$\varepsilon = \pm i^{\frac{n(n-1)}{2}} \left[ g_{11}\, g_{22} \dots g_{nn} \right]^{-1/2}. \tag{1.5.2}$$

In the present case the above reduction only shows that if $I \neq (0)$ is a two-sided ideal $\subset C$, then I contains some element $\alpha + \beta P$, where the scalars $\alpha, \beta$ are not both zero. Now $(\alpha + \beta P)(\alpha - \beta P) = (\alpha - \beta P)(\alpha + \beta P) = \alpha^2 - \beta^2$. Hence if $\alpha^2 \neq \beta^2$, then $\alpha + \beta P$ has an inverse, which implies that $1 \in I$. Hence $I = (1) = C$.

There remains the possibility $\alpha^2 = \beta^2$, that is I contains only such elements $\alpha + \beta P$ which are scalar multiples of either $1 + P$ or $1 - P$. We show that C effectively contains two proper two-sided ideals related to these cases.

Let us define the two elements

$$E_+ = \frac{1+P}{2}, \qquad E_- = \frac{1-P}{2}.$$

We have the relations

$$E_+^2 = E_+, \; E_-^2 = E_-, \; E_+ E_- = E_- E_+ = 0,$$

which follow from (1.5.1). These relations are expressed by saying that $E_+$ and $E_-$ are mutually annihilating idempotents. Both $E_+$ and $E_-$ belong to the center Z of C that is they are central idempotents. Hence if a and b are any elements of C, then

$$(aE_+) \; (bE_-) = (ab) \; (E_+E_-) = 0, \text{ etc.}$$

On the other hand, since

$$E_+ + E_- = 1,$$

if we write for any $a \in C$,

$$a_+ = a E_+, \quad a_- = a E_-,$$

then

$$(1.5.3) \qquad a_+ + a_- = a, \qquad a_+ a_- = 0.$$

The sets $I_+ = \{ a E_+ \}$ and $I_- = \{ a E_- \}$, where a runs over C, form two sided ideals in C, since $E_+$ and $E_-$ belong to Z. An element $p^+ \in C$ belongs to $I_+$ if and only if $E_+ p^+ = p^+$ or equivalently if $E_- p^+ = (1 - E_+) p^+ = 0$. Indeed, if $p^+ \in I_+$, then $p^+ = a\, E_+$, thus $E_+ p^+ = a\, E_+^2 = a\, E_+ = p^+$; and if $E_+ p^+ = p^+$, then by definition $p^+ \in I_+$. Similarly $q^- \in I_-$ if and only if $E_- q^- = q^-$ or equivalently if $E_+ q^- = 0$. Let us prove that the decompositon (1.5.3) of a into an element of $I_+$ and an element of $I_-$ is unique. Indeed if $a = p^+ + q^-$, $p^+ \in I_+$, $q^- \in I_-$, then $E_+ a = E_+ p^+ + E_+ q^- = E_+ p^+ = p^+$ and $E_- a = E_- p + E_- q^- = E_- q^- = q^-$, and thus $p^+ = a_+$, $q^- = a_-$. Summing up, we have proved the following theorem:

*If* $n = \dim C_1$ *is odd and the scalars are the complex numbers, then the Clifford algebra is the direct sum of two mutually annihilating two-sided ideals* $I_+$ and $I_-$. *These ideals are generated by two mutually annihilating central idempotents* $E_+$ *and* $E_-$.

In particular $I_+$ is a subalgebra of C, whose unit element is $E_+$, and $I_-$ is a subalgebra of C, whose unit element is $E_-$. We shall now prove that $I_+$ and $I_-$ are simple. Let us first observe that the elements $e_B E_+$ generate $I_+$. But this system of generators is redundant and $I_+$ is already spanned by those elements $e_A\, E_+ = e_{\alpha_1} e_{\alpha_2} \cdots e_{\alpha_h}\, E_+$, where

$$(1.5.4) \qquad 1 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_h \leqslant n - 1,$$

that is where $e_n$ does not enter as a factor in $e_A$ ($E_+$ itself is included among these elements, corresponding to the case when A is empty ). Indeed if $e_B$ contains $e_n$ as a factor, then $e_B P = \lambda e_A$, where $\lambda \neq 0$ is a scalar and A is the sequence complementary to B with respect to 1,2,..., n. Hence, $e_B P = \lambda e_A$, where $e_A$ does not $e_n$ as a factor. Thus we have $e_B E_+ = e_B P^2 E_+ = (e_B P)(P E_+) = \lambda e_A\, E_+$, from which the assertion follows. Similarly $I_-$ is spanned by the elements $e_A E_-$, where the sequence A satisfies (1.5.4). Since the number of such elements $e_A$ is $2^{n-1}$, it follows that $\{e_A E_+\}$ is a basis of $I_+$ and $\{e_A E_-\}$ is a basis of $I_-$. In particular $I_+$ and $I_-$ both have dimension $2^{n-1}$.

Now let $\tilde{C}$ be the Clifford algebra built on the (n - 1)-dimensional vector space having basis $e_1, e_2, \ldots, e_{n-1}$ and associated with the quadratic form $(x \mid y) = \sum_{i=1}^{n-1} g_{ij} x_i y_j$. Then the mapping $e_A \rightarrow e_A E_+$ defines an isomorphism of $\tilde{C}$ onto $I_+$ and $e_A \rightarrow e_A E_-$ defines an isomorphism of $\tilde{C}$ onto $I_-$. Since n - 1 is even, we know that $\tilde{C}$ is simple, and it follows that $I_+$ and $I_-$ are simple.

We have thus proved that *when* n *is odd*, *the algebra* C *over the complex field of scalars is the direct sum of two simple algebras*. Hence C falls into the category of the so-called *semi-simple* algebras.

If the field of scalars is the real numbers, then we must distinguish between two cases. If the number $\varepsilon$ in (1.5.2) is real, the above reasoning can be carried out without change and we obtain exactly the same conclusions as in the complex case.

If, on the other hand, $\varepsilon$ is imaginary, the algebra C over the real numbers is simple. By what precedes, this follows immediately from the fact that now $1 + \beta e_1 e_2 \ldots e_n$ has an inverse for any $\beta$ since

$$(1 + \beta e_1 e_2 \ldots e_n)(1 - \beta e_1 e_2 \ldots e_n)$$

$$1 - \beta^2 (-1)^{\frac{n(n-1)}{2}} g_{11} g_{22} \cdots g_{nn} = 1 - \beta^2 \varepsilon^2 g_{11}^2 g_{22}^2 \cdots g_{nn}^2$$

is never zero.

1.6. We shall now investigate the connection between the Clifford algebras and the familiar quaternions.

As is well known, the algebra Q of the quarternions over the real field is spanned by the four elements 1, j, k, $\ell$, where 1 is the unit element of Q and the other three elements satisfy the relations

(1.6.1) $\quad j^2 = k^2 = \ell^2 = -1,$

(1.6.2) $\quad jk = -kj = \ell, \; k\ell = -\ell k = j, \; \ell j = -j\ell = k,$

(1.6.3) $\quad jk\ell = -1.$

It is worth while mentioning that any two of these systems of relations implies the third one.

Consider now the real Clifford algebra C associated with the Euclidean metric and corresponding to the case $n = 3$. Then C is of dimension $2^3 = 8$ and its subalgebra $C_+$ spanned by the elements $1, e_1e_2, e_2e_3, e_3e_1$ is of dimension $2^2 = 4$. The correspondences

(1.6.4) $\quad 1 \leftrightarrow 1, \; j \leftrightarrow e_3e_2, \; k \leftrightarrow e_1e_3, \; \ell \leftrightarrow e_2e_1$

establish an isomorphism between Q and C. Indeed $(e_3e_2)^2 = -e_3^2e_2^2 = -1$ and similarly for the other two equations (1.6.1). Furthermore $e_3e_2 \cdot e_1e_3 = e_3^2e_2e_1 = e_2e_1$, $e_1e_3 \cdot e_3e_2 = e_1e_2 = -e_2e_1$ and similarly for the other two relations in (1.6.2).

The real quaternions form a division algebra. Indeed, if for a quaternion

(1.6.5) $\quad q = q_0 + q_1 j + q_2 k + q_3 \ell,$

we define its conjugate by

$$\bar{q} = q_o - q_1 j - q_2 k - q_3 \ell$$

and its norm by

$$N(q) = q\bar{q} = \bar{q}q = q_o^2 + q_1^2 + q_2^2 + q_3^2 ,$$

then $N(q) = 0$ if and only if $q = 0$. Thus every non zero element of Q has an inverse given by $q^{-1} = \bar{q}/N(q)$. By a theorem of Weierstrass and Frobenius all finite dimensional division algebras over the real field are isomorphic to either the complex field or to the quaternion algebra.

Hamilton and his school professed that the quaternions make the study of vectors in three-space unnecessary since every vector can be considered as the vectorial part $q_1 j + q_2 k + q_3 \ell$ of a quaternion (1.6.5). The above isomorphism shows, however, that this interpretation is grossly incorrect since the vectorial part of a quaternion behaves with respect to coordinate transformation like a bivector or "axial" vector and not like an ordinary or "polar" vector.

Let us now pass to the quaternions over the complex field (also called biquaternions). These are of the form (1.6.5), where the coefficients $q_o$, $q_1$, $q_2$, $q_3$ are complex numbers. The complex quaternions do not form a division algebra, since e.g. $(1 + ij)(1 - ij) = 1 - i^2 j^2 = 1 - 1 = 0$; this is in agreement with another theorem of Weierstrass and Frobenius which asserts that there exists no finite dimensional division algebra over the complex field.

The algebra of the complex quaternions (biquaternions) can be considered as an algebra of dimension 8 over the real field spanned by the elements $1, i, j, k, \ell, ij, ik, i\ell$. If we complete (1.6.4) by the correspondence $i \leftrightarrow e_1e_2e_3$, then $ij \leftrightarrow e_1e_2e_3 \cdot e_3e_2 = e_1$, $ik \leftrightarrow e_1e_2e_3 \cdot e_1e_3 = e_2$, $i\ell \leftrightarrow e_1e_2e_3 \cdot e_2e_1 = e_3$. Since $e_1e_2e_3$ belongs to the center and $(e_1e_2e_3)^2 = -1$, we have thus established an isomorphism between the complex quaternions and the real Clifford algebra C corresponding to the case $n = 3$.

Let us observe finally that we can also establish an isomorphism between the real quaternions Q and the real Clifford algebra C for $n = 2$ defined by $g_{11} = g_{22} = -1, g_{12} = g_{21} = 0$, by putting $1 \leftrightarrow 1$, $j \leftrightarrow e_1$, $k \leftrightarrow e_2$, $\ell \leftrightarrow e_1e_2$. Indeed, $(e_1e_2)^2 = -e_1^2e_2^2 = -1$, $e_2 \cdot e_1e_2 = -e_2^2e_1 = e_1$, $e_1e_2 \cdot e_1 = -e_1^2e_2 = e_2$. This isomorphism is of no clearcut geometric interest, since "polar" and "axial" vectors of C appear intermingled.

1.7. The definition of a Clifford algebra is based on the concept of a scalar square or a symmetric scalar product. In this section we shall consider this concept more closely.

Let E be an n-dimensional vector space over the real or the complex field. A scalar product or metric on E is here defined as a function which to every couple (x,y) of vectors of E associates a scalar $(x \mid y)$ with the following properties

$$\text{(I)} \quad (x \mid y) = (y \mid x),$$

$$\text{(II)} \quad (\alpha x_1 + \beta x_2 \mid y) = \alpha (x_1 \mid y) + \beta (x_2 \mid y).$$

If $e_1, e_2, \ldots, e_n$ is any basis of E and $x = \sum_{i=1}^{n} x_i e_i$,

$y = \sum_{i=1}^{n} y_i e_i$, then

$$(1.7.1) \qquad (x \mid y) = \sum_{i=1}^{n} g_{ij}\, x_i\, y_i \text{ with } g_{ij} = (e_i \mid e_j).$$

Conversely, if we are given a basis $e_1, e_2, \ldots, e_n$ and a matrix $(g_{ij})$ with $g_{ij} = g_{ji}$, then (1.7.1) defines a scalar product in E for which $(e_i \mid e_j) = g_{ij}$.

Now we prove that *there exists an orthogonal basis of* E, *i.e. a basis* $f_1 .f_2, \ldots, f_n$ *such that* $(f_i \mid f_j) = 0$ *for* $i \neq j$. Suppose first that $(x \mid y) = 0$ for any couple of vectors x, y in E. In this case any basis is orthogonal (this is the case corresponding to the Grassmann algebra, see p. 7 ). Assume then that there exist two vectors a and b in E such that $(a \mid b) \neq 0$. From this it follows that there exists a vector whose scalar square is not zero. Indeed, if either $(a \mid a) \neq 0$ or $(b \mid b) \neq 0$, then we are finished. If not, then $(a + b \mid a + b) = 2(a \mid b) \neq 0$.

Now let $f_1 \in E$ be a fixed vector such that $(f_1 \mid f_1) \neq 0$. An arbitrary vector $x \in E$ can be decomposed in a unique way into $x = x' + \lambda f_1$ in such a manner that $(x' \mid f_1) = 0$, that is $x'$ is orthogonal to $f_1$. We find $\lambda = (x \mid f_1)/(f_1 \mid f_1)$. Hence E is spanned by $f_1$ and vectors orthogonal to $f_1$. It follows that these vectors form an (n - 1) dimensional subspace F of E. For the space F we have the same alternative as for the space E. In the first case any basis in F together with $f_1$

gives an orthogonal basis of E. In the second case an element $f_2 \in F$ with $(f_2 | f_2) \neq 0$ will permit a reduction to a subspace of dimension $n - 2$, and so on.

In this way we find a basis $f_1, f_2, \ldots, f_n$ such that $(f_i | f_j) = 0$ for $i \neq j$ and $(f_i | f_i) = \mathcal{E}_i \neq 0$ for $i \leqslant r$ $(\leqslant n)$ and $(f_i | f_i) = 0$ for $r+1 \leqslant i \leqslant n$. Writing $x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} x_i' f_i$ we obtain

$$(x | x) = \sum_{i,j=1}^{n} g_{ij}\, x_i\, x_j = \sum_{i=1}^{n} \mathcal{E}_i\, x_i'^2 = \sum_{i=1}^{r} \mathcal{E}_i\, x_i'^2 .$$

It follows from well known theorems about quadratic forms that the number $r$ is independent of the mode of reduction; $r$ is called the rank of the quadratic form and is also equal to the rank of the matrix $(g_{ij})$. In particular $r = n$ if and only if $\det(g_{ij}) \neq 0$. We call the case $r = n$ regular, the case $r < n$ singular; this terminology is consistent with the definitions to be given in Chap. III, section 3.3.

In sections 1.2 and 1.4 we constructed a basis of the Clifford algebra associated with a regular metric form. Again denoting the mutually orthogonal basis vectors constructed above by $e_i$ and setting $e_A = e_{\alpha_1} e_{\alpha_2} \ldots e_{\alpha_p}$ $(1 \leqslant \alpha_1 < \alpha_2 < \cdots < \alpha_p \leqslant n)$ we have shown that these elements form a basis of the Clifford algebra in the regular case. The main difficulty was to prove the linear independence of the $e_A$. As a matter of fact, all these results can be proved in the singular case, but the proof of the independence requires other methods since the elements

$e_A^{-1}$ used in section 1.4 do not necessarily exist. The final proof will be given at the end of section 1.10. One of the main tools is given here since it is closely connected with the results of the present section. Let us also observe that the results of section 1.5 concerning the center, simplicity, etc. do not hold in the singular case. As an exercise we recommend to the reader to determine the center of a Grassmann algebra.

*Any vector space with singular metric can be imbedded in a vector space with regular metric.* Let $E$ be a vector space with the orthogonal basis $e_1, e_2, \ldots, e_n$, that is $(e_i \mid e_k) = 0$ for $i \neq k$, and suppose that

$$(e_i \mid e_i) = g_{ii} \begin{cases} \neq 0, & 1 \leqslant i \leqslant r < n, \\ = 0, & r+1 \leqslant i \leqslant n. \end{cases}$$

Introduce $n - r$ new vectors $e^{r+1}, \ldots, e^n$ and complete the definition of the metric by setting

$$(e^p \mid e^q) = 0, \quad (e_i \mid e^p) \begin{cases} \neq 0, & i = p \\ = 0, & i \neq p \end{cases} \quad (p,q = r+1, \ldots, n).$$

We claim that the space $\overline{E}$ spanned by the $n + (n - r) = 2n - r$ vectors $e_1, \ldots, e_n$; $e^{r+1}, \ldots, e^n$ is regular. To prove this, we show that to every vector $0 \neq \overline{x} \in \overline{E}$ there exists a $\overline{y} \in \overline{E}$ such that $(\overline{x} \mid \overline{y}) \neq 0$, the latter scalar product being defined by distributivity. Let

$$\overline{x} = \lambda_1 e_1 + \ldots + \lambda_r e_r + \mu_{r+1} e_{r+1} + \ldots + \mu_n e_n + \nu_{r+1} e^{r+1} + \ldots + \nu_n e^n .$$

If there is some $\lambda_k \neq 0$, we take $\bar{y} = e_k$, if all the $\lambda_k$ are $= 0$, but some $\mu_p \neq 0$, we take $\bar{y} = e^p$; if all the $\lambda_k$ and $\mu_p$ are $0$ and $\nu_q \neq 0$, we take $\bar{y} = e_q$. In either of these cases $(\bar{x} \mid \bar{y}) \neq 0$, hence the extended metric is regular, as asserted.

1.8. Let $E$ be a vector space of dimension $n$ over the real or the complex field with an arbitrary metric $(x \mid y)$ and $C$ the associated Clifford algebra. Further let $e_1, e_2, \ldots, e_n$ be a basis of $E$. If $x = \sum x_i e_i$ and $y = \sum y_i e_i$ are arbitrary vectors in $C$, we can write

(1.8.1) $$xy = \tfrac{1}{2}(xy + yx) + \tfrac{1}{2}(xy - yx).$$

By (1.1.7), the first term on the right hand side is the scalar product $(x \mid y)$. The second term is

(1.8.2) $$\begin{aligned}\tfrac{1}{2}(xy - yx) &= \tfrac{1}{2} \sum_{i,j} (x_i y_j - x_j y_i)\, e_i e_j \\ &= \tfrac{1}{4} \sum_{i,j} (x_i y_j - x_j y_i)(e_i e_j - e_j e_i)\end{aligned}$$

We know that if $e_1, e_2, \ldots, e_n$ form an orthogonal basis, the distinct $e_i$ anticommute, i.e. $e_i e_j + e_j e_i = 0$ for $i \neq j$. In this case we can also write

(1.8.3) $$\tfrac{1}{2}(xy - yx) = \sum_{i<j} (x_i y_j - x_j y_i)\, e_i e_j.$$

With the symbol

(1.8.4) $$\tfrac{1}{2}(xy - yx) = [x;y]$$

(1.8.1) can be written in the form

(1.8.5) $$xy = (x \mid y) + [x;y].$$

If the vectors $x$ and $y$ are collinear, i.e. $y = \lambda x$, then and only then all $x_i y_j - x_j y_i = 0$, that is $[x;y] = 0$, and we have $xy = (x \mid y)$. If, on the other hand, the vectors $x$ and $y$ are orthogonal, i.e. $(x \mid y) = 0$, then we have $xy = [x;y]$ .

The components in (1.8.2) are clearly skew-symmetric. This suggests that there is a close connection between Clifford numbers and skew-symmetric tensors. We shall devote the subsequent sections to the study of this connection. We begin by recalling some general aspects of tensor algebras

1.9. A tensor algebra is a particular case of a graded algebra. We call an algebra $\mathcal{A}$ with unit element a graded algebra, if it is the direct sum of subspaces $\mathcal{A}_h$ $(h = 0,1,2,\ldots)$, $\mathcal{A} = \sum_{h=0}^{\infty} \mathcal{A}_h$ such that $\mathcal{A}_h \mathcal{A}_k \subset \mathcal{A}_{h+k}$. An element belonging to the subspace $\mathcal{A}_h$ is said to be homogeneous of degree h.

Again let $E$ be an n-dimensional real or complex vector space with basis $e_1, e_2, \ldots, e_n$. The tensor algebra $T(E)$ over $E$ is the infinite dimensional algebra whose basis is formed by all products

$$(1.9.1) \qquad e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_h} ,$$

where $0 \leqslant h < \infty$ and $(i_1, i_2, \ldots, i_h)$ is any set of indices $1,2,\ldots, n$ with possible repetitions. $T(E)$ is the set of all finite linear combinations of elements (1.9.1) with real or complex coefficients, respectively. The product of two elements (1.9.1) is formed by juxtaposition and is extended by distributivity to any pair of elements of $T(E)$. This multiplication is associative but not commutative. Let $T_h(E)$ be the subspace of $T(E)$ spanned by all elements (1.9.1) which contain exactly h factors. Then $T(E)$ is the direct sum of these $T_h(E)$ $(0 \leqslant h < \infty)$ and $T_h(E) \otimes T_k(E) \subset T_{h+k}(E)$, i.e. $T(E)$ is a graded algebra as defined above. The dimension of $T_h(E)$ is $n^h$; the subspace $T_o(E)$ can be identified with the field of scalars, $T_1(E)$ can and will always be identified with E. It is clear that any other basis of E yields the same algebra $T(E)$.

We are interested mostly in skew-symmetric tensors. Let $x_1, x_2, \ldots, x_p$ be p vectors of $E = T_1(E)$. Define the symbol $\varepsilon^{i_1 i_2 \ldots i_p}$ to be 0 if two superscripts are equal and to be $+1$ or $-1$ according as $(i_1, i_2, \ldots, i_p)$ is an even or odd permutation of $(1,2,\ldots,p)$. The commutator $[x_1, x_2, \ldots, x_p]$ of $x_1, x_2, \ldots, x_p$ is defined by

$$(1.9.2)\quad [x_1, x_2, \ldots, x_p] = \sum \varepsilon^{i_1 i_2 \ldots i_p} x_{i_1} \otimes x_{i_2} \otimes \ldots \otimes x_{i_p},$$

where the summation is extended over all permutations $(i_1, i_2, \ldots, i_p)$ of $(1,2,\ldots,p)$. It would be equivalent to ask that all the indices $i_1, \ldots, i_p$ run independently through all values $1, 2, \ldots, p$.

Now let $e_1, e_2, \ldots, e_n$ be a basis of E and let $S_p$ be the subspace of $T_p(E)$ spanned by the commutators

$$(1.9.3)\quad [e_{\alpha_1}, e_{\alpha_2}, \ldots, e_{\alpha_p}]$$

with $1 \leqslant \alpha_1 < \alpha_2 < \ldots < \alpha_p < n$. We remind the reader that according to our convention the "empty" commutator, which stands for 1, is also to be considered. It spans the subspace $S_o$, which will be identified with the field of scalars. Also $S_1 = T_1(E)$ is identified with E. The commutators (1.9.3) are clearly linearly independent since, for example, the product $e_{\alpha_1} \otimes e_{\alpha_2} \otimes \dots \otimes e_{\alpha_p}$ appears only in $[e_{\alpha_1}, e_{\alpha_2}, \dots e_{\alpha_p}]$. Thus $S_p$ is an $\binom{n}{p}$ -dimensional subspace of $T_p(E)$. In particular, for $p > n$ the space $S_p$ reduces to zero.

Let us prove that $S_p$ is independent of the choice of the basis $e_1, e_2, \dots e_n$. To this end let $e'_1, e'_2, \dots e'_n$ be another basis of $E$. Then we have the reversible transformation formulae

$$e'_i = \sum_{j=1}^{n} \gamma_{ij} e_j .$$

For a fixed sequence $1 \leqslant \alpha_1 < \alpha_2 < \dots < \alpha_p \leqslant n$ and a variable sequence $1 \leqslant \beta_1 \leqslant \beta_2 \leqslant \dots \leqslant \beta_p \leqslant n$, and an arbitrary permutation $\beta_{j_1}, \beta_{j_2}, \dots, \beta_{j_p}$ of this sequence we set

$$D(\beta_{j_1}, \dots, \beta_{j_p}) = \begin{vmatrix} \gamma_{\alpha_1 \beta_{j_1}} & \cdots & \gamma_{\alpha_1 \beta_{j_p}} \\ . & . & . \quad . \quad . \\ . & . & . \quad . \quad . \\ \gamma_{\alpha_p \beta_{j_1}} & \cdots & \gamma_{\alpha_p \beta_{j_p}} \end{vmatrix} = \varepsilon^{j_1 \dots j_p} D(\beta_1, \dots, \beta_p) .$$

We observe that these expressions vanish when two $\beta$ are equal.

With this notation and self-explanatory abbreviations for the ranges of the summations, we can write

$$[e'_{\alpha_1}, \dots, e'_{\alpha_p}] = \sum_{(i)} \varepsilon^{i_1 \dots i_p} \sum_{(\beta_j)} \gamma_{\alpha_{i_1} \beta_{j_1}} \cdots \gamma_{\alpha_{i_p} \beta_{j_p}} e_{\beta_{j_1}} \otimes \quad \otimes e_{\beta_{j_p}}$$

(1.9.4)

$$= \sum_{(\beta_j)} D(\beta_{j_1}, \dots, \beta_{j_p}) \, e_{\beta_{j_1}} \otimes \dots \otimes e_{\beta_{j_p}}$$

$$= \sum_{(\beta)} D(\beta_1, \dots, \beta_p) \sum_{(j)} \varepsilon^{j_1 \dots j_p} e_{\beta_{j_1}} \otimes \dots \otimes e_{\beta_{j_p}}$$

$$= \sum_{(\beta)} D(\beta_1, \dots, \beta_p) [e_{\beta_1}, \dots, e_{\beta_p}] .$$

It follows that the elements (1.9.4) are contained in the subspace spanned by the elements (1.9.3) and vice versa, which proves that $S_p$ is independent of the choice of the basis.

An arbitrary element belonging to $S_p$ is called a skew-symmetric tensor of order $p$ or a $p$ -vector. The simplest elements, namely the commutators of $p$ vectors, are called simple $p$ -vectors. We normalize these simple $p$ -vectors by introducing the divided commutators

(1.9.5) $$[a_1; a_2; \dots; a_p] = \frac{1}{p!} [a_1, a_2, \dots, a_p]$$

and call the left hand side the simple $p$ -vector spanned by the vectors $a_1, a_2, \dots, a_p$. A simple $p$-vector is zero if and only if the vectors which span it are linearly dependent. From computation analogous to those just given for the basis elements it follows that any $p$ linearly independent linear combinations of $p$ given vectors span the same simple $p$ -vector up to a scalar factor. Thus a simple $p$-vector is essentially the attribute of the subspace spanned by the vectors in question and not of the individual vectors themselves. Clearly, the sum (or linear combination) of simple $p$ -vectors is a $p$ -vector. The converse follows from the definition of $p$ -vectors.

In particular, the space $S_n$ consists, up to a scalar factor, of a unique (simple) n-vector. An n-vector is also called a pseudo-scalar.

From (1.9.2) we obtain the expansion formulae for commutators, namely

$$(1.9.6)\qquad [x_1,\dots, x_p] = \sum_{k=1}^{p} (-1)^{k-1}\, x_k \otimes [x_1,\dots,\hat{x}_k,\dots x_p]$$

$$= \sum_{k=1}^{p} (-1)^{k+p}\, [x_1,\dots,\hat{x}_k,\dots,x_p] \otimes x_k,$$

where the circumflex indicates that the factor $x_k$ is omitted. For divided commutators these formulae read

$$(1.9.7)\qquad [x_1;\dots;x_p] = \frac{1}{p} \sum_{k=1}^{p} (-1)^{k-1}\, x_k \otimes [x_1;\dots;\hat{x}_k;\dots;x_p]$$

$$= \frac{1}{p} \sum_{k=1}^{p} (-1)^{k+p} [x_1;\dots;\hat{x}_k;\dots;x_p] \otimes x_k.$$

We shall not write out explicitly the formulae (not needed here) which correspond to the the general Laplace expansion formulae of determinants.

We stop this digression concerning the general theory of skew-symmetric tensors here and proceed to applications in Clifford algebras.

1.10. Let C be the Clifford algebra associated with an arbitrary metric form $(x \mid y)$ (cf. 1.1). Given any p elements $x_1, x_2, \dots, x_p$ of E, we again define their commutator by

$$(1.10.1)\qquad [x_1,x_2,\dots, x_p] = \sum \varepsilon^{i_1 \dots i_p}\, x_{i_1} x_{i_2} \dots x_{i_p},$$

where the summation is extended over all permutations $(i_1, i_2, \dots, i_p)$ of $(1,2,\dots,p)$. We also introduce the divided commutators

(1.10.2) $$[x_1; \ldots; x_p] = \frac{1}{p!}[x_1, \ldots, x_p].$$

Let $e_1, \ldots, e_n$ be a basis of E. We show now that the divided commutators of all the sequences $e_{i_1}, e_{i_2}, \ldots, e_{i_p}$ $(p = 0, 1, 2, \ldots, n;\ 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n)$ form a (possibly redundant) basis of C. Later we shall prove that we really have a basis, that the above commutators are linearly independent. Obviously, in the case of mutually orthogonal $e_1, \ldots, e_n$ these commutators are identical to the products $e_A$ considered in sections 1.2-1.6.

The number of factors in a product is called the length of the product. We are going to prove the following assertion. Let $x, \ldots, y, \ldots, z$ be a sequence of vectors in E and $x', \ldots, y', \ldots, z'$ a permutation of this sequence. Form the (Cliffordian) products $\Pi = x \ldots y \ldots z$ and $\Pi' = x' \ldots y' \ldots z'$. Then $\Pi' = \pm \Pi + \Sigma'$, where the sign is + or -, according as the permutation is even or odd and $\Sigma'$ is a linear combination of products of vectors belonging to proper subsequences of the given sequence.

From this and the formulae of definitions (1.10.1) and (1.10.2) it follows immediately that

(1.10.3) $$\Pi = [x; \ldots; y; \ldots; z] + \Sigma,$$

where again $\Sigma$ is a linear combination of products "shorter" than $\Pi$.

The proof is founded on the basic relations

$$(1.10.4) \qquad xy + yx = 2(x \mid y),$$

in particular

$$(1.10.5) \qquad x^2 = xx = (x \mid x).$$

Our assertion follows readily from the fact that the interchange of two adjacent elements changes the sign of the original product and adds to it a scalar multiple of a shorter product, as seen from

$$\begin{aligned} \dots xzyu\dots &= \dots x(2(y \mid z) - yz)u\dots \\ &= -(\dots xyzu\dots) + 2(y \mid z)\,(\dots xu\dots). \end{aligned}$$

This completes the proof of (1.10.3).

If the sequence $x, \dots y, \dots, z$ contains repetitions, the commutator in (1.10.3) is zero. So it follows by iterated use of the above reduction that the product $\Pi$ is equal to a linear combination of products formed without repetitions of elements of proper subsequences of the given sequence.

Now clearly every element of C can be expressed as a linear combination of products of the $e_i$. If we replace the relations (1.10.4) and (1.10.5) by the more detailed ones:

$$(1.10.6) \qquad e_i e_k + e_k e_i = 2(e_i \mid e_k) = 2g_{ik} ,$$

and in particular

$$(1.10.7) \qquad e_i^2 = e_i e_i = (e_i \mid e_i) = g_{ii},$$

it follows from the preceding by induction on the lengths of the products that an arbitrary element of C can be expressed as a linear combination of the divided commutators $[e_{i_1} ; e_{i_2} ; \dots ; e_{i_p}]$ $(p = 0, 1, 2, \dots, n;\ 1 \leq i_1 < i_2 < \dots < i_p \leq n)$. Hence, as asserted, these simple p-vectors form a - possibly redundant- basis of the Clifford algebra.

The analogous p-vectors, considered as elements of the tensor algebra have been proved in section 1.9 to be linearly independent. We are going to prove that this independence subsists in the present case in spite of the relations (1.10.6) and (1.10.7).

First consider the regular case. It has been proved in sections 1.2 and 1.4 that the products $e_A$ form a basis (in the strict sense) of the Clifford algebra. The number of the above divided commutators is the same as that of the products $e_A$ (namely $\sum_{p=0}^{n} \binom{n}{p} = 2^n$), and since they also span the Clifford algebra, it follows that they also form a basis and, in particular, that they are linearly independent.

If the metric of the given vector space E is singular, we imbed the space E in a space $\bar{E}$ provided with a regular metric, as explained at the end of section 1.7. If the dimension of E is n and the metric form of E has rank $r < n$, the space $\bar{E}$ will have dimension $n + (n - r) = 2n - r > n$. Let $\bar{e}_1, \ldots, \bar{e}_n$ be an arbitrary basis of E and complete it to a basis of $\bar{E}$ by adding $n - r$ suitable vectors $\bar{e}_{n+1}, \ldots, \bar{e}_{2n-r}$. Since the space $\bar{E}$ is regular, the p-vectors consisting of the larger basis with subscripts proceeding in the natural order are linearly independent. Hence those, whose subscripts run from 1 to n, a fortiori are linearly independent and consequently form a basis (in the strict sense) of the Clifford algebra associated with the original space E. If we take an orthogonal basis $e_1, \ldots, e_n$ of E, the corresponding p-vectors are identical with the products $e_A$.

Hence, as asserted in section 1.7 ( p. 24 ), these products also form a basis.

1.11 In this section we shall consider briefly some applications of the foregoing concepts to analysis.

We introduce the differential operator nabla

$$\nabla = \sum_{i=1}^{n} e_i \frac{\partial}{\partial x_i},$$

which maps each scalar valued function $f(x)$ into a vector valued function, which is called the gradient of $f(x)$, that is $\nabla f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} e_i = \text{grad } f$. Suppose now that our vectors are imbedded in a Clifford algebra associated with the Euclidean metric. Then it makes sense to speak of the square of the operator $\nabla$ and we get

$$(1.11.1) \qquad \nabla^2 = ( \nabla \mid \nabla ) = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} = \Delta,$$

that is the Laplace operator.

In the relation $ab = (a \mid b) + [a;b]$ (cf. (1.8.5 )) let us replace the vector $a$ by $\nabla$ and the vector $b$ by the vector field $V(x) = \sum_{i=1}^{n} V_i(x)\, e_i$. This yields

$$\nabla V = \sum_{i=1}^{n} \frac{\partial V_i}{\partial x_i} + \sum_{i<j} \left(\frac{\partial V_j}{\partial x_i} - \frac{\partial V_i}{\partial x_j}\right) e_i\, e_j,$$

which can also be written

$$(1.11.2) \qquad \nabla V = \text{div } V + \text{rot } V.$$

Similarly we have

$$V\nabla = \operatorname{div} V - \operatorname{rot} V,$$

where, in $V\nabla$, the operator $\nabla$ acts on V from the right.

If $u + iv = f(x + iy)$ is a complex valued function of the complex variable $x + iy$, the Cauchy-Riemann partial differential equations

$$(1.11.3) \qquad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

can also be written as

$$(1.11.4) \qquad \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(u + iv) = 0.$$

The second equation in (1.11.3) shows that $(u,-v)$ is the gradient of a function h, i.e.

$$(1.11.5) \qquad u = \frac{\partial h}{\partial x}, \quad -v = \frac{\partial h}{\partial y}.$$

Using the first relation in (1.11.3) we obtain

$$\Delta h = \operatorname{div} \operatorname{grad} h = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = 0,$$

i.e. h is a harmonic function. Conversely if h is harmonic, then u and v defined by (1.11.5) satisfy the Cauchy-Riemann equations.

Now let C be the real Clifford algebra spanned by the elements $1, e_1, e_2, e_2 e_1$ and $C_+$ its subalgebra spanned by 1 and $e_2e_1$. Since $(e_2e_1)^2 = -e_2^2 e_1^2 = -1$, this subalgebra $C_+$ is isomorphic to the field of complex numbers. Hence equation (1.11.4) can also be written in the form

$$\left(\frac{\partial}{\partial x} + e_2e_1\frac{\partial}{\partial y}\right)(u + e_2e_1 v) = 0.$$

Inserting $e_2^2 = 1$ between the two factors and using associativity and anticommutativity, we get

$$(e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y})\ (e_1 u - e_2 v) = 0.$$

As we know, this relation holds if and only if $(u,-v)$ is the gradient of a harmonic function. It follows that

$$(e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y})\ (e_1 u + e_2 v) = 0,$$

if and only if $(u,v)$ is the gradient of a harmonic function.

In this form the result can be generalized immediately to any number of dimensions. Let $V$ be a vector field which satisfies the relation

$$(1.11.6) \qquad \nabla V = (\sum e_i \frac{\partial}{\partial x_i})\ (\sum e_i V_i) = 0.$$

By (1.11.2) this means that div $V = 0$ and rot $V = 0$. The second relation implies that $V$ is the gradient of a function $h$, $V = \text{grad}\ h$, while the first relation implies that

$$\Delta h = \text{div grad } h = \text{div } V = 0,$$

i.e. $h$ is a harmonic function. Conversely, if $h$ is a harmonic function, $V = \text{grad}\ h$ satisfies (1.11.6).

However, in contrast to the classical case of functions of a complex variable, this last functional relation is not transitive. The relations $\nabla_x V = 0$, $\nabla_V U = 0$ do not imply the relation $\nabla_x U = 0$. The same remark applies to quaternion functions of quaternion variables, treated at length by Fueter and his pupils.

The $\nabla$ operator can be defined in any regular metric and relations analogous to (1.11.1) and (1.11.2) are also valid. In the Lorentz case the operator $\nabla$ plays a basic role in Maxwell's equations, when written in Clifford language (see section 4.17), and in Dirac's equation (see Chap. VI.).

1.12 The remainder of this chapter is devoted to some important but rather intricate matters, which, however, will not be needed for the understanding of the later chapters.

Let $E$ be an n-dimensional real or complex vector space with a metric $(x \mid y)$ and let $T = T(E)$ be the tensor algebra over $E$ (cf. section 1.9). Let $I$ be the ideal of $T$ generated by the elements

$$x \otimes x - (x \mid x),$$

where $x$ runs through $E = T_1$. This ideal $I$ consists of all finite linear combinations of elements of the form

$$u \otimes (x \otimes x - (x \mid x)) \otimes v ,$$

where $u, v \in T$. We shall prove that the quotient algebra $T/I$ is isomorphic to the Clifford algebra $C$ associated with $E$.

In his book [6] Chevalley defines $C$ as the quotient algebra $T/I$.

The proofs will be based on certain algebraic operations, called derivations, which have their origin in algebraic topology. We treat them at a greater length than needed for our immediate purpose; they will be most useful to us also in other connections. For the following definition of derivation we refer to Chevalley [7].

Let $\mathcal{A} = \sum_{h=0}^{\infty} \mathcal{A}_h$ be a graded algebra ( cf. p. 27 ). We set $\mathcal{A}_h = 0$ for $h < 0$. A map $D$ of $\mathcal{A}$ into $\mathcal{A}$ is called *a derivation of degree* $\nu$ in $\mathcal{A}$ if it satisfies the following conditions.

1) D is linear, i.e. for $u, v \in \mathcal{A}$ we have $D(\lambda u + \mu v) = \lambda Du + \mu Dv$;

2) D is homogeneous of degree $\nu$, i.e. there exists an integer $\nu$ such that $D\mathcal{A}_h \subset \mathcal{A}_{h+\nu}$ for every h;

3) if $u \in \mathcal{A}_h$, $v \in \mathcal{A}$, then

$$(1.12.1) \quad D(uv) = Du.v + (-1)^{h\nu} u.Dv.$$

We can immediately give a more general formulation to (1.12.1). Introduce the *main involution* J (cf. section 1.4) of $\mathcal{A}$ by putting $Ju = u$ or $Ju = -u$ if u is *homogeneous* and of even or odd degree, respectively. Extending J by linearity, we get an *automorphism* of $\mathcal{A}$ with $J^2 =$ the identity, i.e. J is an *involution*. The relation (1.12.1) is then equivalent to

$$(1.12.2) \qquad D(uv) = Du \,.\, v + J^{\nu}u.\, Dv.$$

Setting $u = 1$, $v = 1$, we find: *For every derivation* D *we have* $D1 = 0$.

Let D, D' be two derivations of $\mathcal{A}$ of degrees $\nu$ and $\nu'$ respectively. *Then*

$$(1.12.3) \qquad \Delta = D\, D' - (-1)^{\nu\nu'} D'D$$

*is again a derivation.* The proof is straightforward. It is evident that $\Delta$ is linear and homogeneous of degree $\nu + \nu'$, so it is only necessary to check (1.12.2). Applying the right hand side of (1.12.3) to the product uv and using (1.12.2), the assertion reduces to the two relations

(1.12.4) $J^{\nu} D' - (-1)^{\nu\nu'} D' J^{\nu} = 0$ and $D J^{\nu'} - (-1)^{\nu\nu'} J^{\nu'} D = 0$,

which follow from each other by symmetry. To prove (1.12.4), assume that $u$ is homogeneous of degree $\mu$. Then $Du$ is homogeneous of degree $\mu + \nu$ and

$$J^{\nu'} Du = (-1)^{(\mu+\nu)\nu'} Du = (-1)^{\nu\nu'} D(-1)^{\mu\nu'} u = (-1)^{\nu\nu'} D J^{\nu'} u,$$

which is the second of the relations (1.12.4) written in a slightly altered form.

In particular, *if* $\nu$ *and* $\nu'$ *are both odd*, $D D' + D'D$ *is a derivation, and if* $\nu$ *is odd*, $D^2$ *is a derivation.*

Generally, we shall be mostly concerned with derivations of odd degree and in particular with the cases $\nu = \pm 1$. In the case of an *odd degree* $\nu$, the formulae (1.12.1) and (1.12.2) become

(1.12.1') $D(uv) = Du . v + (-1)^h u . Dv$

for $u \in \mathcal{A}_h, v \in \mathcal{A}$, and

(1.12.2') $D(uv) = Du . v + Ju . Dv$

for $u, v \in \mathcal{A}$. The operator relations (1.12.4) become

(1.12.4') $DJ + JD = 0$

In the case when the degree $\nu$ is *even*, the formula (1.12.2) becomes

(1.12.2") $D(uv) = Du . v + u . Dv$,

for any $u, v \in \mathcal{A}$, recalling the rule for the ordinary derivation (differentiation) of a product. In this case formula (1.12.4) becomes

$$(1.12.4'') \qquad D\,J - JD = 0.$$

1.13. Now let $\mathcal{A}$ be the tensor algebra T considered in section 1.9. For simplicity we shall omit here the sign of multiplication $\otimes$ used there, that is, we shall write uv instead of $u \otimes v$ and, for example, $Ju \cdot Dv$ instead of $Ju \otimes Dv$. As usual, we denote by $x, y, z, x_1, x_2, \ldots$ elements of $T_1 = E$, or, in other words, vectors.

Assume now that a linear mapping $x \longrightarrow Dx$ such that $x \in T_1$ and $Dx \in T_{1+\nu}$ is given for all elements x of $T_1$. We will prove that D can be extended to a derivation (of degree $\nu$) in T in a unique way. We restrict ourselves to the case of an odd $\nu$. The proof in the even case is similar.

Let us first examine some implications of the postulates entering into the definition of a derivation. Here induction means induction on the length of products (cf. p. 32).

a) Setting in (1.12.2') $u = x_1$, $v = x_2 \ldots x_m$, we obtain by induction

$$(1.13.1) \quad D(x_1 x_2 \ldots x_m) = \sum_{k=1}^{m} (-1)^{k-1}\, x_1 \ldots x_{k-1} \cdot Dx_k \cdot x_{k+1} \ldots x_m.$$

Conversely, if u and v are products of certain vectors or linear combinations of such products, and Du, Dv, D(uv) are defined by (1.13.1) and linearity, then (1.12.2') will be satisfied. To show this, write $u = x_1 \ldots x_f$, $v = x_{f+1} \ldots x_m$, and define Du and Dv by (1.13.1). Then $Du \cdot v + Ju \cdot Dv$ yields D(uv) as defined by (1.13.1).

b) If we have a finite sequence $x_1, x_2, \ldots$ of vectors such that all subsequences satisfy relations of the structure of (1.13.1) then, under usual distributivity assumptions, all vectors which are linear combinations of the $x_k$ will do the same. It will be sufficient to prove this for two such vectors $y = \sum_i \alpha_i x_i$ and $z = \sum_k \beta_k x_k$. We have

$$D(yz) = \sum_{i,k} \alpha_i \beta_k D(x_i x_k) = \sum_{i,k} \alpha_i \beta_k (Dx_i \cdot x_k - x_i \cdot Dx_k)$$

$$= \left(\sum_i \alpha_i Dx_i\right)\left(\sum_k \beta_k x_k\right) - \left(\sum_i \alpha_i x_i\right)\left(\sum_k \beta_k Dx_k\right)$$

$$= Dy \cdot z - y \cdot Dz.$$

After these preliminary remarks we are able to prove our statement. First we observe that $DT_0 = 0$, since $D1 = 0$, as shown above. Dx is linear in x, and is determined by its values on the elements $e_k$ of a basis of $T_1$. As we have seen in a), D is to be defined by (1.13.1) for all the products formed of elements of the basis. Thus, as shown in b), products of arbitrary vectors also satisfy (1.13.1). Then, again by a), (1.12.2') will also be generally satisfied. This proves the existence and the uniqueness of the extension of D to T. The formula (1.13.1) and the postulate of linearity yield this extension without reference to any basis.

1.14. As will be seen, the derivations of degree-1 deserve particular attention. According to the general definition given in section 1.12 they map $T_1$ into $T_0$ and this mapping is linear. This being the case, let g(y) be an arbitrary linear function defined for $y \in T_1$. There exists a unique derivation $\delta_g u$ in T such that $\delta_g y = g(y)$ for $y \in T_1$. For the product of the vectors $y_1, y_2, \ldots, y_m$

we have, according to (1.13.1),

$$\delta_g(y_1 \cdots y_m) = \sum_{k=1}^{m} (-1)^{k-1} \; y_1 \cdots y_{k-1} \cdot g(y_k) \cdot y_{k+1} \cdots y_m .$$

Since g(y) is a scalar, this can also be written

$$(1.14.1) \quad \delta_g(y_1 \cdots y_m) = \sum_{k=1}^{m} (-1)^{k-1} \; g(y_k) \cdot y_1 \cdots \hat{y}_k \cdots y_m \; ,$$

where, as before, the circumflex indicates that the factor $y_k$ is omitted. By means of the last formula and the linearity postulate, $\delta_g u$ is given for all $u \in T$. In the present case the multiplication rule (1.12.2') becomes

$$(1.14.2) \quad \delta_g(uv) = \delta_g u . v + Ju . \delta_g v .$$

We shall show that $\delta_g^2 = 0$, i.e. $\delta_g^2 u = \delta_g(\delta_g u) = 0$ for $u \in T$. Since $\delta_g^2$ is a derivation, as proved in section 1.12 (p. 40 ), it is sufficient to show that $\delta_g^2 y = 0$, where $y \in T_1$. But this is obvious, since $\delta_g^2 y = \delta_g(g(y))$ and g(y) is a scalar. If h(y) is another linear function defined for $y \in T_1$, it follows by polarization that $\delta_g \delta_h + \delta_h \delta_g = 0$.

We are most interested in the effect of the operation $\delta_g$ on skew-symmetric tensors. We shall show that $\delta_g$ makes a (p-1)-vector of a p-vector and a simple (p-1)-vector of a simple p-vector. The first assertion is easy to prove. Incidentally, it also follows from the second.

Consider the commutator $[y_1, y_2, \ldots, y_p]$ of order p. We first show that $\delta_g[\;\;]$ is the linear combination of commutators of order p-1. This is equivalent to the first assertion above.

We have, in view of (1.9.2) and (1.14.1),

$$(1.14.3)\quad \begin{aligned}\delta_g[y_1,\ldots,y_p] &= \delta_g \sum_{(\alpha)} \varepsilon^{\alpha_1\ldots\alpha_p} y_{\alpha_1}\ldots y_{\alpha_p}\\ &= \sum_{k=1}^{p} (-1)^{k-1} \sum_{(\alpha)} \varepsilon^{\alpha_1\ldots\alpha_p} g(y_{\alpha_k})\, y_{\alpha_1}\ldots\hat{y}_{\alpha_k}\ldots y_{\alpha_p}.\end{aligned}$$

Set (for a moment) $\alpha_k = \beta$. The p products $y_{\alpha_1}\ldots\hat{y}_\beta\ldots y_{\alpha_p}$, where the $y_{\alpha_i} \neq y_\beta$ appear in a fixed order, but the "empty" symbol $\hat{y}_\beta$ occurs at an arbitrary place, are all identical to $\prod y_{\alpha_i}$, $\alpha_i \neq \beta$, the factors being taken in the fixed order. The term in the double sum (1.14.3) which corresponds to the permutation $\alpha_1,\ldots,\alpha_{j-1},\beta,\ldots,\alpha_p$ can be written

$$(-1)^{j-1}\, \varepsilon^{\alpha_1\ldots\alpha_{j-1}\beta\ldots\alpha_p}\, g(y_\beta) \prod y_{\alpha_i}.$$

Also we have

$$\begin{aligned}(-1)^{j-1}\varepsilon^{\alpha_1\ldots\alpha_{j-1}\beta\ldots\alpha_p} &= \varepsilon^{\beta\alpha_1\ldots\alpha_{j-1}\ldots\alpha_p}\\ &= (-1)^{\beta-1}\varepsilon^{\alpha_1\ldots\alpha_i\ldots\alpha_p},\end{aligned}$$

where the last symbol only has p-1 superscripts $\alpha_i \neq \beta$ in a given order. The contribution of the p terms considered is consequently

$$p(-1)^{\beta-1}\varepsilon^{\alpha_1\ldots\alpha_i\ldots\alpha_p}\, g(y_\beta) \prod y_{\alpha_i}.$$

By this we have proved that

$$\delta_g [y_1, \ldots, y_p]_p$$

(1.14.4)

$$= p \sum_{k=1}^{p} (-1)^{k-1} g(y_k) [y_1, \ldots, \hat{y}_k, \ldots, y_p]_{p-1}$$

which for the divided commutators (cf. (1.9.5)) becomes

$$\delta_g [y_1; \ldots; y_p]_p$$

(1.14.5)

$$= \sum_{k=1}^{p} (-1)^{k-1} g(y_k) [y_1; \ldots; \hat{y}_k; \ldots; y_p]_{p-1}$$

The subscripts attached to the brackets indicate the order of the commutators. Formulae (1.14.4) or (1.14.5) show, among other things, that $\delta_g S_p \subset S_{p-1}$, as asserted above.

Next we are going to prove our second assertion on p. 43 namely that a $\delta$-derivation makes a simple (p-1)-vector $V_{p-1}$ of a simple p-vector $V_p$. We may suppose that the vectors $y_k$ which span $V_p$ are linearly independent. Otherwise $V_p = 0$ and consequently $\delta_g V_p = 0$. Then the vectors $y_k$ span a p-dimensional space $E_p$. Now we distinguish two cases. If all the $g(y_k) = 0$, i.e. g is "orthogonal" to the space $E_p$, then the right hand sides of (1.14.4) or (1.14.5) vanish, i.e. $\delta_g V_p$ is again equal to the zero (p-1)-vector. Hence our assertion is trivial in this case. In the opposite case, there exists a (p-1)-dimensional subspace $E_{p-1}$ of $E_p$ which is orthogonal to g. Indeed, let $e_1, \ldots, e_p$ be a basis of $E_p$ and set $g(e_k) = \alpha_k$. Then all $\alpha_k = 0$, if and only if g is orthogonal to $E_p$. If at least one of them, say $\alpha_p \neq 0$, then g is orthogonal to the p-1 linearly independent vectors $\alpha_p e_k - \alpha_k e_p$ $(k = 1,\ldots,p-1)$. The space spanned by these vectors is the subspace $E_{p-1}$ of $E_p$ which is orthogonal to g. As pointed out in section 1.9, up to a non-vanishing scalar factor $V_p$ is equal to the p-vector spanned by any p linearly independent vectors in $E_p$. If we chose in $E_p$ one arbitrary vector z not in $E_{p-1}$ and p-1 linearly independent vectors $z_1,\ldots, z_{p-1}$ in $E_{p-1}$, we have $g(z) \neq 0$, $g(z_k) = 0$, $k=1,\ldots, p-1$. Hence, according to (1.14.5), up to a non-vanishing factor $\delta_g V_p$ is equal to the (p-1)-vector spanned by $z_1,\ldots, z_{p-1}$.

Summing up: Let $V_p \neq 0$ and $E_p$ the p-dimensional space spanned by $V_p$. If g is orthogonal to $E_p$, we have $\delta_g V_p = 0$. If g is not orthogonal to $E_p$, $\delta_g V_p$ is a simple non zero (p-1)-vector lying in that subspace $E_{p-1}$ of $E_p$ which is orthogonal to g.

1.15. The linear operation we shall introduce in the present section is not a derivation itself, but it is closely connected with a derivation of degree +1.

Let $x \in T_1$ be a fixed vector and $y_1,\ldots,y_m \in T_1$ be arbitrary vectors. We define the operation $M_x$ acting on the product of the latter vectors by the formula

$$M_x (y_1\cdots y_m) = \sum_{k=0}^{m} (-1)^k \; y_1\cdots y_k \, x y_{k+1}\cdots y_m, \tag{1.15.1}$$

or, in another form,

$$(1.15.2) \qquad M_x(y_1 \ldots y_m) = xy_1 \ldots y_m + \sum_{k=1}^{m} (-1)^k \, y_1 \ldots \, y_k x y_{k+1} \ldots y_m .$$

For linear combinations of such products we define $M_x$ by linearity. It will be clear **in a moment that** this does not lead to any contradiction.

The effect of the operator $M_x$ on commutators is of particular interest. It is easy to verify that

$$(1.15.3) \qquad M_x [y_1, \ldots, y_p] = [x, y_1, \ldots, y_p] .$$

Hence, in contrast to $\delta_g$, which lowers the order of a simple p-vector by 1, $M_x$ raises it by 1.

The relation $M_x \, . \, 1 = x \neq 0$ shows that the mapping $u \longrightarrow M_x u$ is not a derivation. However, $(M_x - x).1 = x - x = 0$; in the parenthesis x stands for the operator which maps u into xu. Hence the operator $\overline{M}_x = M_x - x$ might produce a derivation given by $u \longrightarrow (M_x - x)u = M_x u - xu$. This is in fact the case and the degree of this derivation is clearly $+1$. We can use the general scheme developed in detail in section 1.13. It follows immediately from (1.15.2) for products of vectors such as $u = y_1 \ldots y_f$, $v = y_{f+1} \ldots y_m$, and then by linearity, for any elements u, v of the tensor algebra T, that

(1.15.4) $$\bar{M}_x(uv) = \bar{M}_x u.v + J u . \bar{M}_x v.$$

For $M_x$ this gives

(1.15.5) $$M_x(uv) = M_x u.v + J u . (M_x - x)v .$$

We collect the following special formulae:

(1.15.6) $$\begin{cases} M_x . 1 = x, \; M_x y = xy - yx, \; M_x x = 0, \\ \bar{M}_x . 1 = 0, \; \bar{M}_x y = -yx, \; \bar{M}_x x = -x^2. \end{cases}$$

From (1.15.3) we obtain

$$M_x^2 [y_1, y_2, \dots, y_p] = [x, x, y_1, y_2, \dots y_p] = 0.$$

Hence $M_x^2 = 0$, when acting on any skew-symmetric element. However, this is also true when $M_x^2$ acts on any element of the tensor algebra. We shall prove this although only the skew-symmetric case is needed in the sequel.

According to (1.15.5) and (1.15.6), we have for any $u \in T$

$$M_x(xu) = 0 - x(M_x - x)u \quad \text{or} \quad (M_x x + xM_x)u = x^2 u,$$

or the operator equation

$$M_x x + x M_x = x^2.$$

From this it follows that

$$\bar{M}_x^2 = (M_x - x)^2 = M_x^2 - M_x x - x M_x + x^2 = M_x^2 .$$

But, by virtue of (1.15.6) and (1.15.2), for any $y \in T_1$,

$$\bar{M}_x^2 y = \bar{M}_x(-yx) = -yx \, . \, x + y \, . \, xx = 0.$$

Thus, $\bar{M}_x^2$ annihilates $T_1$. Since $\bar{M}_x$ is a derivation of odd degree, $\bar{M}_x^2$ is also a derivation according to section 1.12 (p. 40 ). Thus, according to section 1.13 it also annihilates T. This means that $M_x^2 = \bar{M}_x^2 = 0$, as stated above.

We denote the elements of the homogeneous subspaces $T_h$ by $u_h$ $(h = 0,1,2,\ldots)$ and set

$$P_x u_h = \frac{1}{h+1} M_x u_h. \tag{1.15.7}$$

For the divided commutators (see (1.9.5)) formula (1.15.7) yields

$$P_x [y_1; y_2; \ldots; y_p] = [x; y_1; y_2; \ldots; y_p] . \tag{1.15.8}$$

Clearly, also $P_x^2 u = 0$ for any $u \in T$.

1.16. No metric whatever entered into the considerations of sections 1.12 - 1.15. In the remaining sections of Chapter I we again suppose that the n-dimensional real or complex vector space E is provided with a metric given by a scalar product $(x \mid y)$. Let T be the tensor algebra over E and let x be an arbitrary element of $E = T_1$, fixed until further notice. The $\delta_x$-derivation of T will be defined in the same way as the $\delta_g$-derivation was defined before, except that the linear function g(y) on $T_1$ will be identified with the scalar product $(x \mid y)$. Hence $\delta_x u$ $(u \in T)$

is a derivation of degree -1 in T, determined, according to section 1.13, by linearity and by the relation $\delta_x y = (x \mid y)$, $y \in T_1$.

The following assertions are easy to prove. In the case of a regular metric $(x \mid y)$ the family of the derivations $\delta_x$ is as comprehensive as that of the derivations $\delta_g$. More precisely, to every linear function g(y) there corresponds a unique vector x such that $g(y) = (x \mid y)$. The converse is obvious. In the case of a singular metric $(x \mid y)$ the derivations $\delta_x$ form a proper subset of the derivations $\delta_g$.

Next we consider the operator $P'_x = P_x + \delta_x$. According to sections 1.14 and 1.15, $\delta_x^2 u = 0$ and $P_x^2 \; u = 0$ for all $u \in T$, and hence we have

$$(1.16.1) \qquad P'^2_x = (P_x + \delta_x)^2 = P_x \delta_x + \delta_x P_x$$

in T.

From now on we restrict the domain of the operators $P'_x$ to the space S of the skew-symmetric tensors. First we compute the effect of the operators (1.16.1) on a simple p-vector. According to (1.15.8) and (1.14.5), we have

$$\delta_x P_x [y_1;\ldots;;y_p] = \delta_x [x;y_1;\ldots;\; y_p]$$

$$= (x \mid x) [y_1;\ldots;y_p] + \sum_{k=1}^{p} (-1)^k (x \mid y_k) [x;y_1;\ldots;\hat{y}_k;\ldots;y_p];$$

and on the other hand

$$P_x \delta_x [y_1;\ldots;y_p] = P_x \Big( \sum_{k=1}^{p} (-1)^{k-1} (x \mid y_k) [y_1;\ldots;\hat{y}_k;\ldots;\; y_p] \Big)$$

$$= \sum_{k=1}^{p} (-1)^{k-1} (x \mid y_k) \; [x; y_1; \ldots; \hat{y}_k; \ldots; y_p].$$

Hence

$$(1.16.2) \qquad (\delta_x P_x + P_x \delta_x) [y_1; \ldots; y_p] = (x \mid x) \; [y_1; \ldots; y_p].$$

This shows that the operator $\delta_x P_x + P_x \delta_x - (x \mid x)$ or, what is the same, the operator $P'^2_x - (x \mid x)$ annihilates the subspace S of skew-symmetric tensors. Briefly

$$(1.16.3) \qquad P'^2_x - (x \mid x) = 0 \quad \text{or} \quad P'^2_x = (x \mid x),$$

when acting on S. From this it follows by polarization for $x, y \in T_1$, that

$$(1.16.4) \qquad P'_x P'_y + P'_y P'_x - 2(x \mid y) = 0 \quad \text{or} \quad P'_x P'_y + P'_y P'_x = 2(x \mid y),$$

when acting on S.

Next we form the operator $[P'_{x_1}; \ldots; P'_{x_p}]$, that is the divided commutator of the operators $P'_{x_k}$, $k = 1, 2, \ldots, p$ and prove that

$$(1.16.5) \qquad [P'_{x_1}; \ldots; P'_{x_p}] \; 1 = [x_1; \ldots; x_p].$$

The proof proceeds by induction. The relation is correct for $p = 1$, since, in view of the relation $\delta_x 1 = 0$ and of formulae (1.15.6), (1.15.7), $P'_x\ 1 = x$. We assume now that (1.16.5) is correct if p is replaced by p - 1. Then according to (1.9.7) we find that

$$[P'_{x_1};\ldots;P'_{x_p}]\ 1 = \frac{1}{p}\Big(\sum_{k=1}^{p}(-1)^{k-1}P'_{x_k}[P'_{x_1};\ldots;\hat{P}'_{x_k};\ldots;P'_{x_p}]\Big)\ 1$$

$$= \frac{1}{p}\sum_{k=1}^{p}(-1)^{k-1}P'_{x_k}[x_1;\ldots;\hat{x}_k;\ldots;x_p]\ .$$

If we had $P_{x_k}$ instead of $P'_{x_k}$, (1.16.5) would follow from (1.15.8). Therefore, since $P'_x = P_x + \delta_x$, we only have to prove that

$$(1.16.6)\qquad \sum_{k=1}^{p}(-1)^{k-1}\delta_{x_k}[x_1;\ldots;\hat{x}_k;\ldots;x_p] = 0.$$

We find

$$(1.16.7)\qquad (-1)^{k-1}\delta_{x_k}[x_1;\ldots;\hat{x}_k;\ldots;x_p] = \sum_{\substack{i\\ i\neq k}}(-1)^{k-1+i'-1}(x_k|x_i)[\cdot]_{ik},$$

where $[\cdot]_{ik}$ is the divided commutator of the given vectors $x_1,\ldots, x_p$ after the ommision of the vectors $x_k$ and $x_i$ from the sequence, and $i'$ in the exponent is equal to $i$ if $i < k$ and to $i - 1$ if $i > k$. Hence

$$(1.16.8)\qquad (-1)^{k+i'-2} = \begin{cases}(-1)^{k+i}, & i < k,\\ (-1)^{k+i-1}, & i > k.\end{cases}$$

The term $(x_k|x_i)[\cdot]_{ik}$ occurs only in (1.16.7) and in

$(-1)^{i-1}\delta_{x_i}[x_1;\dots;\hat{x}_i;\dots;x_p]$ . The formula (1.16.8) shows that it occurs with opposite signs. This proves (1.16.6) and consequently (1.16.5).

From (1.16.5) we conclude the following fact which is of utmost importance for the sequel.

The operators $[P'_{e_{\alpha_1}};\dots;P'_{e_{\alpha_p}}]$ $(p=0,1,\dots,n;$ $1\leqslant\alpha_1<\alpha_2<\dots<\alpha_p\leqslant n)$ corresponding to a basis $e_1, e_2,\dots, e_n$ of E are linearly independent.

Indeed, a linear dependence would by virtue of (1.16.5), imply the linear dependence of the p-vectors $[e_{\alpha_1};\dots;e_{\alpha_p}]$ of the tensor algebra T. However, the independence of these p-vectors was proved in an entirely trivial way in section 1.9, p. 29 ; cf. also the comments in section 1.10, p. 34.

The main results of this section consist of the equivalent operator relations (1.16.3) and (1.16.4) and of the linear independence of the operators $[P'_{e_{\alpha_1}};\dots;P'_{e_{\alpha_p}}]$. The domain as well as the range of the operation is the space S of the skew-symmetric tensors.

From the results just stated it follows that the algebra generated by the operators $P'_x$ is a Clifford algebra associated with the given metric form $(x\mid y)$, and with the $2^n$ elements $[P'_{e_{\alpha_1}};\dots;P'_{e_{\alpha_p}}]$ as basis, where $e_1,\dots,e_n$ is an arbitrary basis of the given space E.

Indeed, by the method of reduction given in section 1.10, any element of the algebra in question can be expressed as a linear combination of the above p-vectors. They are also linearly independent and hence they form a basis of the algebra.

We point out that in the course of this investigation no mention has been made of the possibly regular or singular character of the metric $(x \mid y)$.

1.17. The (two-sided) ideal I and the quotient algebra T/I, briefly mentioned in section 1.12 will be considered in greater detail in the present section.

Let E be a real or complex n-dimensional vector space provided with a metric $(x \mid y)$ and let T be the tensor algebra associated with E. Form all the differences $x \otimes x - (x \mid x)$, as x runs through $T_1 = E$, form the ideal I these elements generate and the quotient algebra T/I. The elements of T/I are the residue classes of T modulo I. The classes $R_x$, $R_1$, $R_o$ containing the respective elements x,1,0, clearly satisfy the (Cliffordian) equation $R_x^2 - (x \mid x)R_1 = R_o$ (=I). Following Chevalley, we call here T/I The Clifford algebra associated with the metric $(x \mid y)$, the residue classes then being considered as Clifford numbers. A striking advantage of this approach is that the associative law of multiplication automatically follows from that of the tensor algebra.

The main problem is to construct a basis of T/I. We proceeed as in section 1.10, but we have to replace the equations (1.10.4) - (1.10.7) by the congruences

(1.17.1) $$x \otimes y + y \otimes x \equiv 2(x \mid y), \quad x \otimes x \equiv (x \mid x)$$

and/or

(1.17.2) $$e_i \otimes e_k + e_k \otimes e_i \equiv 2g_{ik}\ ,\quad e_i \otimes e_i \equiv g_{ii}\ ,$$

all the congruences being taken modulo I.

Equalities such as (1.10.3) deduced in section 1.10 must also be replaced by similar congruences. By analogy to the final result of section 1.10, we have here the following conclusion.

Any element of the tensor algebra T is congruent modulo I to some skew-symmetric element of T. Hence every residue class contains at least one skew-symmetric element. Moreover T/I admits those residue classes which contain the simple p-vectors $[e_{\alpha_1};\ e_{\alpha_2};\dots;\ e_{\alpha_p}]$, $\alpha_1 < \alpha_2 \dots < \alpha_p$, spanned by elements $e_j$ of a basis of T, as a - possibly redundant - basis.

The next problem is to prove that this basis is not redundant. A direct proof seems to be rather difficult. We shall say more about this in a moment. Here we give a proof based upon two properties of the operators $P'_x$ , established in section 1.16. We are going to prove the following assertion, from which everything follows. Every residue class contains exactly one skew-symmetric element of T.

We recall that $P'_x$ is a linear operator depending on the vector $x \in (T_1 =)E$ and acting on the elements of the subspace S of the skew-symmetric elements of T. The properties in question are the following:

(1.17.3) $$P'^2_x - (x \mid x) = 0,$$

and

(1.17.4) $$[P'_{x_1};\dots;P'_{x_p}]\,1 = [x_1;\dots;x_p]\,.$$

Now let $u \in T$ be arbitrary and let $u \longrightarrow P'_u$ be the natural mapping of T onto the P' - algebra, that is the algebra generated by the operators $P'_x$. From (1.17.3) we infer that, if $v \in I$, then $P'_v = 0$. On the other hand, by (1.17.4) we find that, if $0 \neq s \in S$, then $P'_s\ 1 = s \neq 0$ and hence $P'_s \neq 0$. This shows that I does not contain any non zero skew-symmetric elements. Since, as just shown, every residue class contains at least one skew-symmetric element, it contains exactly one.

We have constructed three different models for the Clifford algebra C associated with a vector space E provided with a metric $(x \mid y)$. First, C was obtained in a formal way, by means of the multiplication rules $x^2 = (x \mid x)$ or $xy + yx = 2(x \mid y)$, where $x, y \in E$, or $e_i e_k + e_k e_i = 2(e_i \mid e_k)$, where the $e_i$ form a basis of E. Secondly, C was obtained as the operator algebra generated by the operators $P'_x$ (p. 53 ). Thus C appeared as a subalgebra of the endomorphisms of the space S of skew-symmetric tensors over E. Thirdly, C appeared as the quotient algebra T/I, where T was the tensor algebra over E and I the ideal of T generated by the elements $x \otimes x - (x \mid x)$. It follows from our discussions that these three models are isomorphic.

Now we indicate the difficulty arising in the "direct" investigation of the ideal I and the quotient algebra T/I. We consider a two-dimensional vector space with the basis elements a and b and the (Euclidean) metric

$$(a \mid a) = (b \mid b) = 1, \quad (a \mid b) = 0.$$

The elements of the associated tensor algebra T will be denoted by u, u', v, v', w, w', ... . The ideal I consists of all finite sums of products of the threetypes

$u \otimes \{a \otimes a - 1\} \otimes u', v \otimes \{b \otimes b - 1\} \otimes v', w \otimes \{a \otimes b + b \otimes a\} \otimes w'$.

It should be shown that I does not contain any skew-symmetric element of T. I can not even show (in a direct way) that the ideal does not contain the element 1, and consequently the entire tensor algebra T.

The difficulty arises from the fact that T and I are infinite dimensional. As shown above, this difficulty can be overcome by the mapping of T onto a suitable finite dimensional algebra.

The method exposed in section 1.16 and in the present section is essentially due to Chevalley [6,pp. 38-42]. However, there are some slight modifications to be mentioned. There is, to begin with, a vector space with a given metric $(x \mid y)$, or $B_o(x,y)$ in Chevalley's notation. We wish to construct a Clifford algebra associated with this metric. According to Chevalley's excellent idea, this can be done by the construction of an operator algebra exhibiting the properties of the desired Clifford algebra. Chevalley uses the operators $\delta_x$, $L_x$, $L'_x = L_x + \delta_x$, acting on the elements of the exterior algebra (here usually called Grassmann algebra)associated with the vector space. The operator $L_x$ designates left multiplication by the vector x in the sense of the exterior algebra. By means of the properties of this algebra it is shown that the $L'_x$

generate a Clifford algebra corresponding to the metric $B\ (x,y)$.

However, an exterior algebra is itself a Clifford algebra (although much easier to handle than a general one). For this reason it seems more appropriate to construct, as done here, an operator algebra acting on a space such as that of the skew-symmetric tensors, which has no Cliffordian character whatever. The operators $P_x$ and $P'_x$ used in section 1.16 and in the present section correspond to the operators $L_x$ and $L'_x$ used by Chevalley. Obviously the two kinds of operators are isomorphic.

Another difference lies in the choice of the basis of the operator algebra. The skew-symmetric basis elements we use have (in the case when the underlying field is of characteristic 0) considerably simpler properties than the products used by Chevalley; cf. in particular formula (1.16.5). As a matter of fact $[L'_{x_1};\ldots;L'_{x_p}]\ 1$ is equal to the exterior product of the vectors $x_1,\ldots,x_p$.

1.18. In view of certain applications, we extend the concepts of the operations $\delta_g$ and $P_x$, previously defined for tensor algebras, to Clifford algebras. We must be cautious, since a Clifford algebra is not a graded algebra (see below). However, it will turn out that no caution is needed as far as $\delta_g$ is concerned.

Incidentally, a Clifford algebra exhibits a certain kind of gradation. An algebra $\mathcal{A}$ is called semi-graded if it is the direct sum of two subalgebras $\mathcal{A}_+$ and $\mathcal{A}_-$ such that

$$\mathcal{A}_+\mathcal{A}_+ \subset \mathcal{A}_+, \quad \mathcal{A}_+\mathcal{A}_- \subset \mathcal{A}_-, \quad \mathcal{A}_-\mathcal{A}_+ \subset \mathcal{A}_-, \quad \mathcal{A}_-\mathcal{A}_- \subset \mathcal{A}_+.$$

The facts contained in formulae (1.4.8) can now be expressed by saying that C is semi-graded. But C is not graded, since $C_p\ C_q$ is not contained in $C_{p+q}$. However the weaker but important relations

(1.18.1) $$C_p C_q \subset \sum_k C_{p+q-2k}, \quad 0 \leqslant k \leqslant \min\,(p,q),$$

hold.

Because of the basic relations (1.10.4) - (1.10.7), any element of a Clifford algebra can be written in infinitely many ways. For example

$$yz = 2(y \mid z) - zy = 2(y \mid z) - zy + y^2 - (y \mid y).$$

These elements, identical in the Clifford algebra associated with the metric $(x \mid y)$, are entirely different when considered as belonging to the tensor algebra. However, applying to them an operation $\delta_g$, as defined in tensorial algebra, we get identical elements of the Clifford algebra.

This fact is best explained by means of the ideal I, treated in the preceding section. We prove that $a \in I$ implies $\delta_g a \in I$. Indeed, if $y \in T_1$, then, according to (1.14.2), $\delta_g(y \otimes y) = g(y)y - y \,.\, g(y) = 0$, since $g(y)$ is a scalar. Since $\delta_g 1 = 0$, we have

(1.18.2) $$\delta_g(y \otimes y - (y \mid y)) = 0.$$

(From this it follows by polarization that $\delta_g(y \otimes z + z \otimes y - 2(y|z)) = 0$.) On the other hand, it follows from (1.14.2) for $u, v, w \in T$ and $\delta_g v = 0$, that

$$\delta_g(u \otimes v \otimes w) = \delta_g u \otimes v \otimes w + Ju \otimes Jv \otimes \delta_g w. \tag{1.18.3}$$

Every element of I is a linear combination of triple products of the above form, with $v$ equal to some $y \otimes y - (y \mid y)$. Thus $\delta_g v = 0$, according to (1.18.2), and $Jv = v$. Consequently, the right hand side of (1.18.3) again belongs to I and our assertion follows. This implies that the operation $\delta_g$ is well defined on the quotient algebra $T/I = C$.

This being the case, the operation $\delta_g$ can be applied to a Clifford number according to the rules given in tensorial algebra, quite independently of the form in which this number may appear. We also have $\delta_g C_p \subset C_{p-1}$.

However, this favorable situation does not exist for the operators $P_x$. A straightforward calculation shows that for $y \in T_1$ we have $P_x(y \otimes y - (y \mid y)) \equiv -\frac{2}{3}(x \mid y)\, y \pmod{I}$ which is $\neq 0$ and does not even belong to I.

A satisfactory solution is offered by the fact that any Clifford number can be expressed in an essentially unique way as linear combinations of commutators of Clifford vectors. (This means that all the possible expressions are equal also when considered as elements of the tensor algebra T.) The operations $P_x$ defined for these expressions according to tensor rules give a uniquely determined result. For example

$$P_x(yz) = P_x((y \mid z) + [y;z]) = (y \mid z)x + [x;y;z]$$

and

$$P_x(zy) = P_x((y \mid z) - [y;z]) = (y \mid z)x - [x;y;z] .$$

Formal addition yields $P_x(yz + zy) = 2(y \mid z)x$, which is correct, since $yz + zy = 2(y \mid z)$. Formal subtraction yields $P_x[y,z] = 2[x;y;z]$, which also is correct, since it is equivalent to $P_x[y;z] = [x;y;z]$. We shall not inquire further into these questions, since we will apply $P_x$ only to linear combinations of p-vectors (also for different p).

Let C be a Clifford algebra, x a vector in C and $V_p$ a simple p-vector in C. We are interested in the product $x V_p$. It follows from (1.18.1) that $x V_p \in C_{p-1} + C_{p+1}$. We wish to show that $x V_p = V_{p-1} + V_{p+1}$, where $V_{p-1}$ is a simple (p-1)-vector and $V_{p+1}$ a simple (p+1)-vector. (similarly for right multiplication by x.) The uniqueness of such a decomposition is obvious since C is the direct sum of the $C_p$.

We first express $V_{p-1}$ and $V_{p+1}$ in a way where the operations $\delta_x$ and $P_x$ are not involved. We write

(1.18.4) $V_{p-1} = \frac{1}{2}(xV_p - (-1)^p V_p x)$ and $V_{p+1} = \frac{1}{2}(xV_p + (-1)^p V_p x)$.

This obviously gives $xV_p = V_{p-1} + V_{p+1}$ (and $(-1)^p V_p x = V_{p+1} - V_{p-1}$). We prove that the right hand side of the first equation (1.18.4) yields a simple (p-1)-vector $V_{p-1}$ orthogonal to x and that the right hand side of the second equation yields a simple (p+1)-vector $V_{p+1}$ spanned by x and $V_p$. The proof will partially follow a proof given in section 1.14.

Consider an orthogonal basis for the space $E_p$ spanned by $V_p$, as indicated in section 1.7. Then $V_p$ can be written as the product of the normalized basis vectors $V_p = f_1 f_2 \dots f_p$. If x is orthogonal to $E_p$, that is to all vectors $f_k$, then x anticommutes with the $f_k$, which gives $xV_p = (-1)^p V_p x$. Hence the right hand side of the first equation of (1.18.4) vanishes and the right hand side of the second of (1.18.4) becomes $xf_1 \dots f_p$, that is the $(p+1)$-vector spanned by x and $V_p$.

If x is not orthogonal to $E_p$, there exists a $(p-1)$-dimensional subspace $E_{p-1}$ of $E_p$ to which x is orthogonal (cf. section 1.14, p. 45). We use this fact in the proof and discussion of the first of the formulae (1.18.4).

Let $V_p = [y_1; \dots; y_p]$ and let $z_1, \dots z_p$ be a normalized basis for the space E spanned by the vectors $y_1, \dots, y_p$ such that $[z_1; \dots; z_p] = [y_1; \dots; y_p]$, and moreover such that the vectors $z_1, \dots, z_{p-1}$ lie in the subspace $E_{p-1}$, and hence are orthogonal to x. The p-vector $V_p = [z_1; z_2; \dots; z_p]$ is, up to a factor $1/p!$, the sum of the expressions

$$\pm A_\alpha = \pm z_{\alpha_1} \dots z_{\alpha_k} z_p \dots z_{\alpha_{p-1}},$$

where the $\alpha_i$ are a permutation of the indices $1, \dots, p-1$. Since x anticommutes with the $z_{\alpha_i}$, we have

$$\tfrac{1}{2}(xA_\alpha - (-1)^p A_\alpha x) = \pm z_{\alpha_1} \dots z_{\alpha_k} B_\alpha \dots z_{\alpha_{p-1}},$$

$$B_\alpha = \tfrac{1}{2}((-1)^k \; xz_p - (-1)^{p+(p-1)-k} z_p x)$$

$$= (-1)^k \; \tfrac{1}{2}(xz_p + z_p x) = (-1)^k \; (x \mid z_p).$$

Since $(x \mid z_p)$ is a scalar, it follows that

$$\pm \tfrac{1}{2}(xA_\alpha - (-1)^p A_\alpha x) = (-1)^k \; (x|z_p)( z_{\alpha_1} z_{\alpha_2} \dots z_{\alpha_{p-1}} ).$$

The ± designates $\varepsilon^{\alpha_1 \dots \alpha_k \, p \dots \alpha_{p-1}} = (-1)^{p-1-k} \varepsilon^{\alpha_1 \dots \alpha_{p-1} \, p}$ $= (-1)^{p-1-k} \varepsilon^{\alpha_1 \dots \alpha_{p-1}}$; hence

$$\tfrac{1}{2}(x V_p - (-1)^p \; V_p x) = \frac{1}{2p!} \sum_\alpha \pm (xA_\alpha - (-1)^p A_\alpha x)$$

$$= (-1)^{p-1} \frac{(x \mid z_p)}{p!} \; p \sum \varepsilon^{\alpha_1 \dots \alpha_{p-1}} z_{\alpha_1} \dots z_{\alpha_{p-1}}$$

(for the factor p cf. the deduction of formula (1.14.4)).

Consequently

$$\tfrac{1}{2}(x V_p - (-1)^p \; V_p x) = (-1)^{p-1} (x \mid z_p)[z_1; \dots ; z_{p-1}] = \delta_x V_p. \tag{1.18.5}$$

For the last identification see (1.14.5) and (1.9.5).

The second of the formulae (1.18.4) becomes in explicit form

$$\tfrac{1}{2} \left\{ x [ y_1; \dots ; y_p ] + (-1)^p [ y_1; \dots ; y_p ] \; x \right\}$$

(1.18.6)

$$= [x; y_1; \dots ; y_p] = P_x [ y_1; \dots ; y_p ] .$$

We shall deduce it from the following assertion to which it is virtually equivalent. The expression

$$(-1)^{k-1} \{ x_k [x_1, \dots , \hat{x}_k, \dots , x_q] + (-1)^{q-1} [x_1, \dots , \hat{x}_k, \dots , x_q] \, x_k \} \tag{1.18.7}$$

is independent of k.

This assertion becomes $x_1x_2 - x_2x_1 = -(x_2x_1 - x_1x_2)$ for $q=2$, which is obvious. We suppose that the assertion corresponding to q-1 is valid, that is

$$(1.18.8)\quad (-1)^{k-1}\{z_k[z_1,\dots,\hat{z}_k,\dots z_{q-1}]+(-1)^{q-2}[z_1,\dots,\hat{z}_k,\dots,z_{q-1}]\,z_k\}$$

is independent of k. According to (1.9.6) we have

$$(1.18.9)\ [x_1,\dots,\hat{x}_k,\dots,x_q]=\tfrac{1}{2}\sum_{\substack{i\\ i\neq k}}((-1)^{i'-1}x_i[\cdot]_{ik})+\tfrac{1}{2}\sum_{\substack{i\\ i=k}}((-1)^{i'+q-1}[\cdot]_{ik}x_i),$$

where, as in formula (1.16.7), the symbol $[\cdot]_{ik}$ denotes the commutator of the elements $x_1,\dots, x_q$ after the omission of $x_k$ and $x_i$ from the sequence, and $i' = i$ if $i<k$ and $i'=i-1$ if $i > k$. By virtue of the assumption relative to (1.18.8), the terms on the right hand side of (1.18.9) corresponding to different values of i are all equal, and consequently equal to the term corresponding to any fixed $j \neq k$. Hence, for brevity setting $[x_1,\dots,\hat{x}_k,\dots,x_q]=[\dots\hat{x}_k\dots]$ and $[\cdot]_{jk} = [\cdot]$, we have

$$[\dots\hat{x}_k\dots] = \tfrac{1}{2}(q-1)\ \{ (-1)^{j'-1} x_j [\cdot]+(-1)^{j'+q-1} [\cdot]\ x_j \}.$$

From this it follows, after obvious reductions modulo 2 in the exponents, that

$$(-1)^{k-1}\{x_k[\dots\hat{x}_k\dots] + (-1)^{q-1}[\dots\hat{x}_k\dots]\ x_k\}$$

$$=\tfrac{1}{2}(q-1)(-1)^{k+j'}\{x_kx_j[\cdot]-[\cdot]x_jx_k+(-1)^q x_k[\cdot]x_j+(-1)^{q-1}x_j[\cdot]\,x_k\}.$$

Defining $k'=k$ if $k < j$ and $k'=k-1$ if $k > j$, we find by symmetry that

$$(-1)^{j-1}\{x_j[\ldots\hat{x}_j\ldots] + (-1)^{q-1}[\ldots\hat{x}_j\ldots]\,x_j\}$$

$$= \tfrac{1}{2}(q-1)\,(-1)^{j+k'}\{x_jx_k[\cdot] - [\cdot]\,x_kx_j + (-1)^q x_j[\cdot]x_k + (-1)^{q-1}x_k[\cdot]\,x_j\}.$$

We assert that the difference of these two expressions, corresponding to k and j respectively, vanishes. Since $(k+j') + (j+k') = 2(k+j) - 1$, $k+j'$ and $j+k'$ have different parity (cf.(1.16.8)). Hence, after suppression of an unessential numerical factor, the difference is

$$(x_k\,x_j + x_j\,x_k)[\cdot] - [\cdot]\,(x_j\,x_k + x_k\,x_j).$$

The last difference vanishes, since $(x_kx_j + x_jx_k) = 2(x_k \mid x_j)$ is a scalar. (As a matter of fact, this is the only place where the Cliffordian character of the vectors enters into this rather involved computation.)

The fact that (1.18.7) is independent of k, which now is proved, yields by means of (1.9.6),

$$(1.18.10)\quad [x_1,\ldots,x_q] = (-1)^{k-1}\tfrac{1}{2}q\{x_k[x_1,\ldots,\hat{x}_k,\ldots,x_q] + (-1)^{q-1}[x_1,\ldots,\hat{x}_k,\ldots,x_q]x_k\}$$

which (in somewhat altered notations) is identical to the formula (1.18.6) to be proved originally.

It is remarkable that in Clifford algebras the expansion formulae (1.9.6) or (1.9.7) can be replaced by the much simpler formulae (1.18.10) or (1.18.6), which show that the successive commutators $[x_1,\ldots, x_i]$, $i = 1,2,\ldots,$ have a very simple recursion law.

We return to the derivations $\delta_x$. From (1.18.5) it follows by linearity that for an arbitrary $u \in C$ we have

(1.18.11) $$\delta_x u = \tfrac{1}{2}(xu - Ju \cdot x).$$

In fact, we can take this as a formula *defining* the derivations $\delta_x$ in Clifford algebras. Let us in any case check the properties of such a $\delta_x$. From (1.18.5) it follows that $\delta_x$ is a linear mapping of $C_p$ into $C_{p-1}$, i.e. it is of degree -1. Moreover if $y \in C_1$, then $\delta_x y = \frac{1}{2}(xy + yx) = (x \backslash y)$, as it should be. A straightforward computation shows that the basic multiplication rule (1.14.2) is also satisfied. However, we want to show this in a more general connection.

Let $\mathcal{A}$ be an arbitrary algebra and $\omega$ an arbitrary automorphism of $\mathcal{A}$. Let a be a fixed element of $\mathcal{A}$ and set

(1.18.12) $$D_a u = \tfrac{1}{2}(au - \omega u . a), \qquad u \in \mathcal{A}$$

(the factor $\frac{1}{2}$ could be omitted or replaced by any other scalar factor). We have for $u, v \in \mathcal{A}$

(1.18.13) $$D_a(uv) = D_a u \cdot v + \omega u \cdot D_a v.$$

Indeed

$$D_a(uv) = \tfrac{1}{2}(a \cdot uv - \omega(uv) \cdot a) = \tfrac{1}{2}(a \cdot uv - \omega u . \omega v \cdot a)$$

and

$$D_a u \cdot v + \omega u \cdot D_a v = \tfrac{1}{2}\left\{(au - \omega u \cdot a)v + \omega u(av - \omega v.a)\right\}$$

$$= \tfrac{1}{2}(auv - \omega u \cdot \omega v \cdot a).$$

This praves (1.18.13), which includes our last assertion concerning $\delta_x$.

From (1.18.13) it follows that $D_a 1 = 0$. However this follows directly from (1.18.12), since $\omega 1 = 1$.

If the algebra $\mathcal{A}$ is graded and the element a homogeneous of degree $\nu$, then $D_a$ is a sort of a derivation of degree $\nu$.

No such simple multiplication rule can be given for the operation $u \to \frac{1}{2}(au + \omega u.a)$ suggested by the operator $P_x$. But one should not forget that $P_x$ itself has no simple multiplication rule either; it is $M_x$ or $\overline{M}_x$ which have one.

However, if, in a Clifford algebra, we define

$$P_x u = \tfrac{1}{2}(xu + Ju \,.\, x), \tag{1.18.14}$$

for $x \in C_1$, then, for p-vectors $u_p$, we again have the expression (1.18.6), that is

$$P_x u_p = \tfrac{1}{2}(xu_p + (-1)^p \; u_p \, x), \tag{1.18.15}$$

originally given for simple p-vectors $V_p$.

1.19. Finally, we give an application of the derivations $\delta_g$ to a problem concerning the square of a bivector. From the fact that a simple p-vector $V_p$ can be written as the product of p mutually orthogonal vectors, it follows by anticommutativity that $V_p^2$ is a scalar. For the case of a simple bivector we give an alternative proof not using orthogonalization. A simple bivector can be written $F = \frac{1}{2}(xy - yx) = [x;y]$, where x and y are vectors. We write (cf.(1.8.1))

$$xy = \tfrac{1}{2}(xy + yx) + \tfrac{1}{2}(xy - yx) = (x \mid y) + [x;y]$$

and

$$yx = (x \mid y) - [x;y].$$

Hence

$$xyyx = ((x \mid y) + [x;y])\,(x \mid y) - [x;y]),$$

or "the identity of Lagrange"

$$(x \mid y)^2 - (x \mid x)(y \mid y) = [x;y]^2,$$

which again proves that the square of a simple bivector is a scalar. There is also a converse statement. We will prove: A necessary and sufficient condition for a bivector F to be simple is that $F^2$ is a scalar. The unproved part of this assertion can be established in the following way.

Let g and h be linear functions on $C_1$, and set $\delta_g F = u$, $\delta_h F = v$. Then $u, v \in C_1$ that is u and v are Clifford vectors. Since, according to our assumption, $F^2$ is a scalar, we have $\delta_g F^2 = 0$, or, by virtue of (1.14.2), $\delta_g F \,.\, F + JF \,.\, \delta_g F = 0$, or, on account of $\delta_g F = u$ and $JF = F$,

$$uF + Fu = 0. \tag{1.19.1}$$

Hence also $\delta_h (uF + Fu) = 0$. According to $\delta_h u = h(u)$, $\delta_h(F) = v$, $Ju = -u$ and formula (1.14.2), we obtain

$$h(u)F - uv + vu + F\,h(u) = 0,$$

or, since h(u) is a scalar,

$$2h(u)F = uv - vu.$$

We assert that, if $F \neq 0$, there exist such linear functions g and h that $h(u) \neq 0$. This being admitted for a moment, we can write

$$F = \frac{1}{2h(u)} (uv - vu), \tag{1.19.2}$$

where the right hand side is a simple bivector, which is the desired result.

Let us look at h(u). The general relation $\delta_h \delta_g + \delta_g \delta_h = 0$

(cf. section 1.14, p. 43) applied to F yields the skew symmetry relation $h(u) + g(v) = 0$, or $h(u) = -g(v)$, which establishes the symmetry in g and h of the formula (1.19.2). Now we prove that, if $F \neq 0$, two linear functions g and h can be found so that $h(u) = -g(v) = \delta_h \delta_g F \neq 0$.

For any fixed basis $e_1, \ldots, e_n$ of $C_1 = E$ we set $g(e_i) = g_i$, $h(e_k) = h_k$, where the scalar values $g_i$ and $h_k$, which can be chosen arbitrarily, determine the linear functions g and h (cf. section 1.13). We find according to (1.14.2) that

$$(1.19.3)\quad \delta_g(e_i e_k) = g_i e_k - g_i e_k, \quad \delta_h(e_i e_k) = h_i e_k - h_k e_i.$$

On the other hand (cf. (1.8.2) and also (4.11.5)) the bivector F can be written

$$(1.19.4)\qquad F = \tfrac{1}{2} \sum_{i,k} F^{ik} \, e_i e_k \qquad F^{ik} = -F^{ki}.$$

Using the last condition and (1.19.3), we easily find that

$$(1.19.5)\qquad u = \delta_g F = \sum_{i,k} F^{ik} \, g_i e_k \,, \quad v = \delta_h F = \sum_{i,k} F^{ik} \, h_i e_k$$

and

$$(1.19.6)\qquad h(u) = -g(v) = \delta_h \delta_g F = \sum_{i,k} F^{ik} \, g_i \, h_k .$$

The assumption $F \neq 0$ is the same as the assumption that not all the coefficients $F^{ik}$ vanish. Under this assumption the skew-symmetric form (1.19.6) does not vanish identically. (If, for example, $F^{12} \neq 0$, we can chose $g_1 = h_2 = 1$ and all the other $g_i$ and $h_k = 0$.) This completes the proof.

We wish to explain, why we have been working with general derivations $\delta_g$ and $\delta_h$ and not with special ones $\delta_x$ and $\delta_y$. In the latter case, one can, in view of (1.18.5), write explicity.

$$u = \delta_x F = \tfrac{1}{2}(xF - Fx), \qquad v = \delta_y F = \tfrac{1}{2}(yF - Fy)$$

and

$$uF + Fu = \tfrac{1}{2}(xF^2 - FxF + FxF - F^2x) = \tfrac{1}{2}(xF^2 - F^2x) = 0$$

if $F^2$ is a scalar. Hence

$$0 = y(uF + Fu) + (uF + Fu)y = y\,Fu + uFy + yuF + Fuy$$
$$= (yF - Fy)u + u(Fy - y\,F) + (yu + uy)\,F + F(uy + yu)$$
$$= 2(vu - uv) + 4(y \mid u)\,F.$$

This is seemingly the same result as above. The difference is that in the case of a singular metric there may not exist two vectors $x, y \in C_1$ such that the scalar product $(y \mid u) \neq 0$. For example, in the case of a Grassmann algebra all scalar products vanish. On the contrary the derivations $\delta_g$ are not influenced by the metric.

To our result about a simple bivector we add the following remark. If $F$ is an arbitrary bivector, then $F^2$ is the sum of a scalar and a 4-vector, that is $F^2 = c_0 + c_4$, where $c_k \in C_k$. Indeed, a priori, $F^2$ could be the sum of a scalar, a bivector and a 4-vector, $F^2 = c_0 + c_2 + c_4$ (cf.(1.18.1)). Performing a reversion (cf. section 1.4, p. 13 ) we obtain $\widetilde{F^2} = \tilde{c}_0 + \tilde{c}_2 + \tilde{c}_4$. From $\widetilde{F^2} = F^2$, $\tilde{c}_0 = c_0$, $\tilde{c}_4 = c_4$, $\tilde{c}_2 = -c_2$, it follows that $c_2 = 0$, which was to be proved. Hence, our result about a simple bivector can be interpreted by saying that in the square of a simple bivector the 4-vector does not appear and vice versa.

## Chapter II

## Rotations and Reflections

2.1 We begin with the following elementary theorem due to Hamilton. *Let* ABC *be a triangle in the plane and suppose that its three angles* $\alpha$, $\beta$, $\gamma$ *are oriented in the same way. Let* P *be an arbitrary point of the plane. If we rotate* P *first around* A *by the angle* $2\alpha$, *then around* B *by the angle* $2\beta$ *and finally around* C *by the angle* $2\gamma$, *then* P *returns to its original position.*

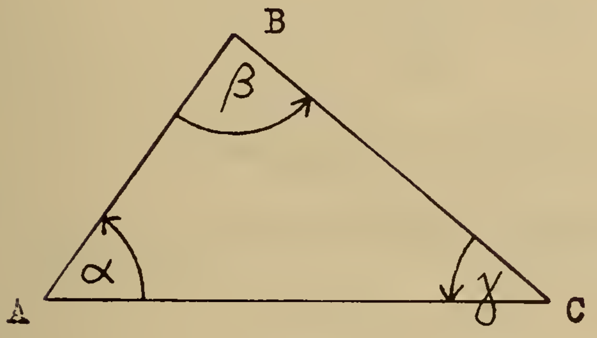


Fig. 1

The proof of this theorem depends on the following principle, also due to Hamilton. *A rotation in the plane around a center* O *by an angle* $2\alpha$ *is equivalent to two successive reflections in two lines of the plane through* O *making an angle* $\alpha$. Indeed, let a and b be two such lines. Let P be any point in the plane. Denote by P' the point obtained by reflecting P in a and by $P''$ the point obtained by reflecting $P'$ in b. clearly $OP = OP' = OP''$ and $\sphericalangle POP'' = 2\alpha$ and thus Hamilton's principle is proved.

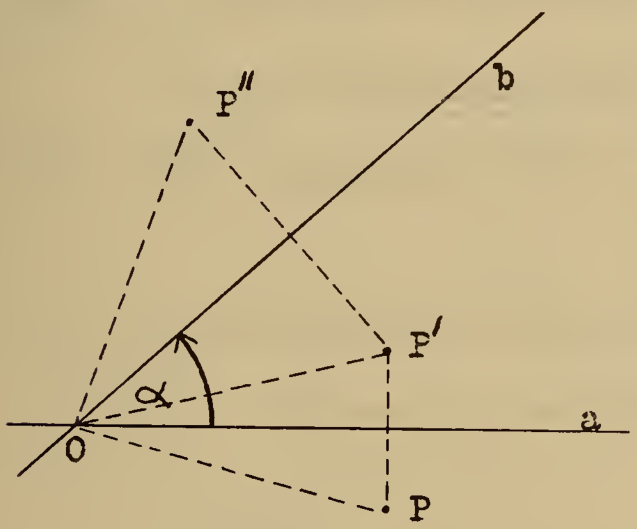


Fig. 2

The theorem stated at the outset follows now by the inspection of Fig. 3. The three rotations can be decomposed into six reflections. The successive positions of the point are P, P', P",P',P"',P',P. Let us observe that the same theorem holds for triangles on the sphere. In both cases the theorem can be extended to an arbitrary polygonal.

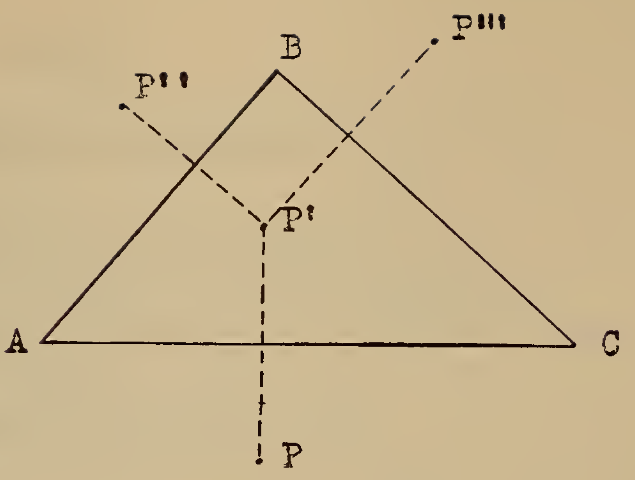


Fig. 3

Let us now consider, in 3-dimensional Euclidean space, a vector $x = OP$ which does not lie in the plane spanned by the lines a, b. We show that a *rotation around an axis through* 0 *and perpendicular to* a *and* b *can also be decomposed into two successive reflections in the lines* a *and* b *respectively*. Indeed, x can be decomposed $x = x_{\parallel} + x_{\perp}$, where $x_{\parallel}$ lies in the plane ab and $x_{\perp}$ is orthogonal to it. Now $x_{\perp}$ is left invariant by the rotation, whereas a reflection changes $x_{\perp}$ into $-x_{\perp}$, thus two successive reflections also leave $x_{\perp}$ invariant. On the other hand the two successive reflections transform $x_{\parallel}$ into the corresponding component of the rotated vector, as follows from the plane case considered above. This proves our assertion.

In the following pages we shall generalize this situation and obtain an analytic expression for rotations in n-space.

2.2 Let E be an n-dimensional vector space over a field F, which in this chapter will be *exclusively* the field R of the *real* numbers. Let there be given on E a *regular* metric $(x \mid y)$. We recall (see sec. 1.7 and 3.3) that such a metric is a function of the two variables $x, y \in E$, taking its value in F and satisfying the following conditions:

(a) $(x \mid y) = (y \mid x)$,

(b) $(\lambda x_1 + \mu x_2 \mid y) = \lambda (x_1 \mid y) + \mu (x_2 \mid y), \quad \lambda, \mu \in F$,

(c) for every $0 \neq x \in E$ there exists a $y \in E$ such that $(x \mid y) \neq 0$.

We also know that if $e_1, e_2, \dots, e_n$ is a basis of E, then for $x = \sum_{i=1}^{n} x_i e_i$ and $y = \sum_{i=1}^{n} y_i e_i$ we have

$$(x \mid y) = \sum_{i,j=1}^{n} g_{ij} \; x_i y_j ,$$

where $g_{ij} = g_{ji}$ and $\det (g_{ij}) \neq 0$. Furthermore there exists a basis of E such that $g_{ij} = 0$ for $i \neq j$, which implies of course that $g_{ii} \neq 0$ $(i = 1, \dots, n)$. From our point of view the two most important particular cases are the Euclidean metric, which, with an appropriate choice of the basis, can be written in the form

$$(x \mid y) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n ,$$

and the Lorentz metric, which can be written in the form

$$(x \mid y) = x_1 y_1 - x_2 y_2 - \dots - x_n y_n .$$

Finally let us recall that two vectors x and y are said to be orthogonal if $(x \mid y) = 0$ and that a non zero vector x is said to be isotropic if $(x \mid x) = 0$.

Let A be an endomorphism of E, i.e. a linear mapping of E into itself. We say that A is an isometry or an automorphism of the metric $(x \mid y)$ if we have

(2.2.1) $$(Ax \mid Ay) = (x \mid y)$$

for all $x, y \in E$. Clearly the totality of the isometries of a given metric forms a group.

Let $V$ be a hyperplane, i.e. a maximal proper subspace in $E$ and suppose that $V$ is regular. This means that no vector of $V$ is orthogonal to $V$ or, in other words, that the restriction of the metric $(x \mid y)$ to $V$ is regular. Then $E$ is the direct sum of $V$ and a straight line $a$ orthogonal to $V$ (see sec. 1.7). Let $x$ be an arbitrary vector of $E$. Then $x$ can be decomposed in a unique way into

$$x = x_{\parallel} + x_{\perp},$$

where $x_{\parallel} \in V$ and $x_{\perp}$ is collinear with $a$. Define the linear operator $S_V$ by setting

$$S_V\, x = x_{\parallel} - x_{\perp} .$$

$S_V$ is obviously linear and is called the *symmetry* in the hyperplane $V$. Since $(x_{\parallel} \mid y_{\perp}) = (x_{\perp} \mid y_{\parallel}) = 0$, we have

$$(S_V x \mid S_V y) = (x_{\parallel} - x_{\perp} \mid y_{\parallel} - y_{\perp}) =$$
$$= (x_{\parallel} \mid y_{\parallel}) + (x_{\perp} \mid y_{\perp}) = (x_{\parallel} + x_{\perp} \mid y_{\parallel} + y_{\perp}) = (x \mid y).$$

Thus the linear transformation $S_V$ is an isometry.

The following theorem is due to E. Cartan and has been extended by Dieudonné to the case of an arbitrary field of scalars:

*Let* $A$ *be an isometry of* $E$. *Then* $A$ *can be written as the product of at most* $n$ *symmetries in regular hyperplanes of* $E$.

The reader can find the proof for an arbitrary field and an arbitrary metric in Artin's book [1 , pp. 129-130]. Here we restrict ourselves to the real field and metrics with Euclidean or Lorentz signature. In the first case the proof is a very simple, due to the absence of singular subspaces. In the second case the proof is less simple, and the reader should postopone its study until he becomes familiar with the content of section 3.6.

We begin by the Euclidean case. The proof proceeds by induction. For $n = 1$ the theorem is obviously true. Let us suppose that the theorem is true for all isometries of an $(n - 1)$ - dimensional Euclidean space.

1) Suppose first that there exists a non zero vector $x \in E$ such that $Ax = x$. Let $H$ be the hyperplane formed by all vectors orthogonal to $x$. The isometry $A$ maps $H$ into itself. Indeed, if $y \in H$, then $(x \mid y) = 0$ and thus $(x \mid Ay) = (Ax \mid Ay) = (x \mid y) = 0$, i.e. $Ay \in H$.

Let $A_H$ be the restriction of $A$ to $H$. It is clear that $A_H$ is an isometry of $H$. By the induction hypothesis there exist $r \leqslant n-1$ hyperplanes $\widetilde{V}_1, \dots, \widetilde{V}_r$ of dimension $(n-2)$ in $H$ such that $A_H$ is the product of the symmetries $S_{\widetilde{V}_1}, \dots S_{\widetilde{V}_r}$ in these hyperplanes.

Call $L$ the one-dimensional subspace of $E$ generated by $x$ and let $V_i$ be the direct sum of $L$ and $\widetilde{V}_i$ $(i = 1, \dots r)$. Clearly $V_i$ is an $(n-1)$-dimensional hyperplane of $E$. It is also clear that the restriction of the symmetry $S_{V_i}$ to $H$ is exactly $S_{\widetilde{V}_i}$ $(i = 1, \dots, r)$ and that $A$ is the product of the $r \leqslant n-1$ symmetries $S_{V_1}, \dots S_{V_r}$.

2) Let now $A$ be an arbitrary isometry (different from the identity) and let $x$ be a vector such that $Ax \neq x$. Since $(Ax \mid Ax) = (x \mid x)$, we have

(2.2.2) $$(Ax - x \mid Ax + x) = 0.$$

Let H be the hyperplane orthogonal to the vector $Ax - x$ and let $S_H$ be the symmetry in H. By virtue of (2.2.2) the vector $Ax + x$ belongs to H and thus

(2.2.3) $$S_H\,(Ax + x) = Ax + x.$$

On the other hand

(2.2.4) $$S_H\,(Ax - x) = x - Ax.$$

Adding relations (2.2.3) and (2.2.4) we obtain. $S_H Ax = x$. Now $S_H A$ is clearly an isometry and hence it follows from part 1) of the proof that $S_H A = S_{V_1} \dots S_{V_r}$, where $r \leq n-1$ and the $S_{V_i}$ are symmetries in hyperplanes $V_i$ $(i = 1,\dots, r)$. Hence we obtain

$$A = S_H\, S_{V_1} \dots S_{V_r}$$

and since $r + 1 \leq n$, the theorem is proved in the case of a Euclidean metric.

We now turn to the proof of the theorem in the case of a Lorentz metric. The proof follows closely the steps of the Euclidean case. We suppose that the theorem is true for all (n-1)-dimensional Lorentz spaces and we shall also make use of the fact that it is true in Euclidean spaces.

1) Suppose first that there exists a non isotropic vector $x \in E$ such that $Ax = x$. The hyperplane H orthogonal to x is either time-like or space-like. In either case the restriction $A_H$ of A to H is the product of the symmetries $S_{\tilde{V}_1}, \dots, S_{\tilde{V}_r}$, where $r \leq n-1$ and $\tilde{V}_1, \dots, \tilde{V}_r$ are regular hyperplanes in H.

Defining the hyperplanes $V_i$ $(i = 1,\dots r)$ of E as the direct sum of $\tilde{V}_i$ and of the line generated by x, we see that $A = S_{V_1} \ldots S_{V_r}$ and that the $V_i$ are regular hyperplanes in E.

2) Let x be any time-like vector and set $Ax = y$. We suppose that $y - x \neq 0$ and we show that $y - x$ is not isotropic. Indeed, since $(y \mid y) = (x \mid x)$, we have

$$(y - x \mid y + x) = 0 \tag{2.2.5}$$

and

$$(y \pm x \mid y \pm x) = (y \mid y) + (x \mid x) \pm 2(y \mid x) \tag{2.2.6}$$

$$= 2 \{ (x \mid x) \pm (y \mid x) \} .$$

Since x is time-like, y is also time-like and (see sec. 3.6) $(x \mid y) \neq 0$. If $(y \mid x) > 0$, then by (2.2.6) $y + x$ is time-like and by (2.2.5) $y - x$ is space-like (see sec. 3.6). If $(y \mid x) < 0$, then by (2.2.6) $y - x$ is time-like. In either case $y - x$ is non isotropic.

From here on the proof is exactly as in the Euclidean case. Let H be the hyperplane orthogonal to $y - x$. Then H is regular, because $y - x$ is non isotropic. Hence $S_H Ax = x$ and, since x is time-like, it follows from part 1) of the proof that $S_H A = S_{V_1} \ldots S_{V_r}$, i.e. $A = S_H S_{V_1} \ldots S_{V_r}$, where $r + 1 \leq n$ and the $V_i$ are regular hyperplanes in E. Thus the theorem is proved also in the case of a Lorentz metric.

2.3 We shall now express the results of the previous section with the help of Clifford numbers. As usual we identify (see sec. 1.4) the vector space E with the subspace $C_1$ of the Clifford algebra C over E associated with the regular metric $(x \mid y)$.

Let y be a fixed non isotropic vector in $E = C_1$. Such a vector y is invertible, and its inverse is $y^{-1} = y/(y \mid y)$.

We want to interpret the transformation.

$$S_y x = y^{-1}xy \,, \qquad x \in E, \tag{2.3.1}$$

where the product $y^{-1}xy$ is to be understood in the sense of the Clifford algebra. The vector $x$ can be decomposed in a unique way into $x = x_1 + x_2$, where $x_1$ is collinear with $y$ and $x_2$ orthogonal to $y$. From $x_1 = \lambda y$ it follows that $x_1$ commutes with $y$, i.e. $y^{-1}x_1y = y^{-1}yx_1 = x_1$. On the other hand $x_2$ anticommutes with $y$ since $x_2y + yx_2 = 2(x_2 \mid y) = 0$ and thus we have $y^{-1}x_2y = - y^{-1}yx_2 = - x_2$. Combining these two results we obtain that $y^{-1}xy = x_1 - x_2$. In other words, $S_y x = y^{-1}xy$ is the vector obtained by reflecting the vector $x$ in the vector $y$. Let $V$ be the hyperplane orthogonal to $y$. Then $x_1$ is the component of $x$ orthogonal to $V$ and $x_2$ the component of $x$ lying in $v$, that is, with our former notations, $x_1 = x_{\perp}$, $x_2 = x_{\parallel}$. Hence we have

$$S_V x = x_{\parallel} - x_{\perp} = - y^{-1}xy. \tag{2.3.2}$$

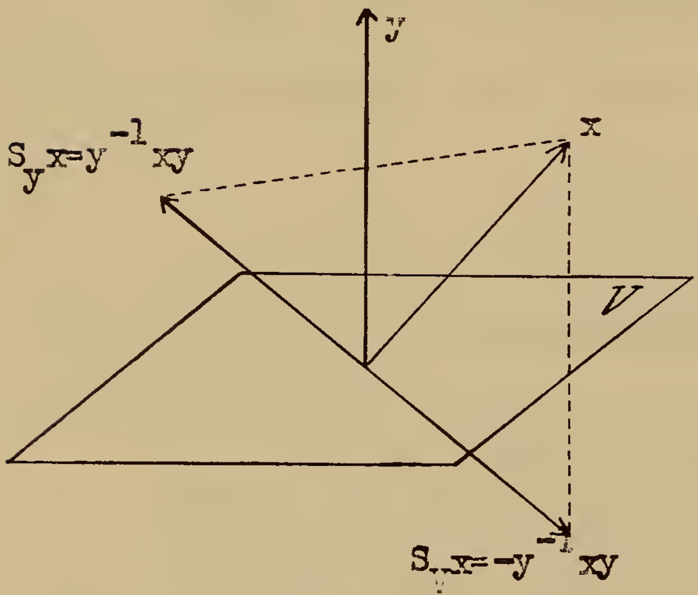


Fig. 4

It follows from the theorem of section 2.2 that *every isometry* A *of* E *can be written in the form*

$$Ax = (-1)^r y_r^{-1} \ldots y_1^{-1}xy_1 \ldots y_r \,, \tag{2.3.3}$$

where $r \leq n$ and $y_1, \ldots y_r$ *are non isotropic vectors in* E.

The vectors $y_1, \ldots, y_r$, where $r$ in general attains its maximal value $r = n$, can be chosen with a very great freedom, so they obviously are not uniquely determined. Even when they are normalized by the requirements $(y_k \mid y_k) = \pm 1$, the formula (2.3.3) still includes in general $n(n-1)$ parameters. On the other hand, an isometry is determined by half this number, i.e. by $\frac{1}{2}n(n-1)$ parameters. Indeed, if $G = (g_{ij})$ is the matrix of the metric, then the fact that A is an isometry is expressed by the matrix equation $A^{\dagger} GA = G$, where $A^{\dagger}$ is the transpose of A. Since G is symmetric, this gives $\frac{1}{2}n(n+1)$ equations for the $n^2$ elements of A. Thus these elements depend on $n^2-\frac{1}{2}n(n+1)=\frac{1}{2}n(n-1)$ parameters.

One of the principal aims of chapters III and IV is to replace (2.3.3) by canonical expressions displaying the correct number of parameters. Our immediate purpose here is to examine to what extent the product $y_1 \ldots y_r$ itself is determined.

2.4 We speak of a rotation or a reflection according as r is even or odd, and note incidentally that the determinant of A is $+1$ when A is a rotation and $-1$ when A is a reflection. We set $y_1 \ldots y_r = s$ and write (2.3.3.) in the form $Ax = \pm s^{-1}xs$. We now prove the following. If s and t are invertible elements of the Clifford algebra C and the simultaneous equations $Ax = \pm s^{-1}xs$ and $Ax = \pm t^{-1}xt$ hold with identical signs, then $t = \xi s$, where $\xi$ belongs to the center of C. Indeed, the two equations give $x(ts^{-1}) = (ts^{-1})x$. Hence $ts^{-1}$ commutes with every vector and thus commutes with every element of C. This means that $ts^{-1} = \xi$ belongs to the center of C. Our assertion $t = \xi s$ is proved.

From now on we suppose that the dimension n of the vector space E is even. Then the center of C is the field of the scalars.

$C_o$ (see sec. 1.5). Thus $t = \lambda s$, where $\lambda$ is a scalar. If, in particular, $s$ and $t$ are products of vectors normalized by the requirement $(y_k | y_k) = \pm 1$, then $t = \pm s$, but there is no way of getting rid of the indeterminacy $\pm 1$. It is inherent to the problem (see sec. 4.12). This is also obvious from the fact that $y_k \to -y_k$ for a fixed $k$ changes the product $s$ into $-s$ but does not change $Ax$.

As in section 1.4, we denote the element obtained from the element $c$ by reversion by $\tilde{c}$. Then we always have $s^{-1} = \lambda \tilde{s}$, where $\lambda$ is a scalar, and $s^{-1} = \pm \tilde{s}$ if the vectors $y_k$ are normalized.

As a further restriction we suppose that $r$ is even also, that is, we restrict ourselves to rotations. In this case $Ax = s^{-1}xs$. Since every Clifford number is the linear combination of products of vectors, the map $x \to s^{-1}xs$ can be extended to an inner automorphism of the Clifford algebra given by $c \to s^{-1}cs$ for any Clifford number $c$.

Here we come in contact with a notion called the Clifford group by Chevalley in his book [6, p. 49]. This group $\Gamma$ is the multiplicative group of all invertible Clifford numbers $s$ such that for every $x \in E = C_1$ the element $s^{-1}xs$ belongs to $C_1$, in other words, that the inner automorphism $c \to s^{-1}cs$ transforms vectors into vectors.

It is worth while to observe that from the property $s \in \Gamma$ alone it follows that $x \to x^1 = s^{-1}xs$ is an isometry. Indeed,

$$(x' | x') = x'^2 = (s^{-1}xs)(s^{-1}xs)$$
$$= s^{-1}x^2s = s^{-1}(x | x)s = (x | x).$$

For more details about the Clifford group and its relation to the spin representation of the group of isometries we refer to Chevalley's book [6].

2.5 Consider now a plane (two dimensional vector space) with the Euclidean metric and let a and b be two orthogonal unit vectors in it. Thus $a^2 = 1$, $b^2 = 1$, $ab+ba = 0$. The plane will be oriented by assuming that a rotation of $+\frac{\pi}{2}$ carries a into b. The Clifford bivector $I = ab$ is independent of the particular choice of a and b. Indeed if a' and b' are two other orthogonal unit vectors with the same orientation, then

$$a' = \cos\theta . a + \sin\theta . b, \qquad b' = -\sin\theta . a + \cos\theta . b$$

and

$$a'b' = (\cos^2\theta + \sin^2\theta)ab + (\cos\theta\sin\theta)(b^2 - a^2) = ab,$$

since a and b anticommute and $a^2 = b^2 = 1$. A change of orientation replaces I by - I. We have $I^2 = (ab)(ab) = -abba = -1$. Thus I shares the basic property of the imaginary unit $i = \sqrt{-1}$, but in contrast to i, it anticommutes with every vector in the plane:

$$xI = -Ix.$$

Indeed, I anticommutes with a and b, and thus it anticommutes with every linear combination of a and b.

If x and $\tilde{x}$ are two vectors oriented in the same way as a and b, such that $|x| = |\tilde{x}| \neq 0$ and $(x | \tilde{x}) = 0$, then $xI = \tilde{x}$ and consequently $Ix = -\tilde{x}$. Indeed, by what we have seen,

$$I = \frac{x}{|x|} \frac{\tilde{x}}{|x|}$$

and thus

$$xI = x\frac{x}{|x|}\,\frac{\tilde{x}}{|x|} = \frac{(x|x)}{|x|\,2}\,.\,\tilde{x} = \tilde{x}.$$

Let us denote by $\text{rot}_\theta$ the operator which rotates every vector of the plane by the angle $\theta$. Then we have

$$\text{rot}_\theta x = \cos\theta . x + \sin\theta . \tilde{x}.$$

This can also be written as

(2.5.1) $$\text{rot}_\theta x = \cos\theta . x + \text{Sin}\theta x I$$

$$= x(\cos\theta + \sin\theta . I) = xe^{\theta I},$$

where $e^{\theta I}$ can be defined either by $\cos\theta + \sin\theta . I$ or by $\sum_{k=0}^{\infty} \frac{(\theta I)^k}{k!}$ .

We also have

$$\cos\theta + \sin\theta . I = (\cos\frac{\theta}{2} + \sin\frac{\theta}{2} . I)^2$$

and thus (2.5.1) can be written as

$$\text{rot}_\theta x = x(\cos\theta + \sin\theta . I)$$

$$= x(\cos\frac{\theta}{2} + \sin\frac{\theta}{2} . I)^2$$

$$= (\cos\frac{\theta}{2} - \sin\frac{\theta}{2} . I)\, x\, (\cos\frac{\theta}{2} + \sin\frac{\theta}{2} . I)$$

$$= e^{-\frac{\theta}{2} I}\, x\, e^{\frac{\theta}{2} I} .$$

The right hand side of the formula

(2.5.2) $$\text{rot}_\theta x = e^{-\frac{\theta}{2} I}\, x\, e^{\frac{\theta}{2} I}$$

is again of the form $s^{-1}xs$. Since $\theta$ is determined modulo $2\pi$, the factor $\frac{\theta}{2}$ is determined only modulo $\pi$ and $e^{\frac{\theta}{2} I}$ is determined only up to a factor ± 1 .

As an exercise, the reader should deduce the important formula (2.5.2) from the Hamilton decomposition of a rotation into two reflections and from the reflection formula (2.3.1)

Hint: If $c^2 = 1$ and $c$ forms the angle $\frac{\theta}{2}$ with $a$, then $c = \cos\frac{\theta}{2} . a + \sin\frac{\theta}{2} . b$ .

Let now $E$ be an n-dimensional Euclidean space and $y$ a vector orthogonal to the plane spanned by the two mutually orthogonal unit vectors $a$ and $b$. Then we set, by definition, $\mathrm{rot}_\theta y = y$. The formual (2.5.2) gives the same result. Indeed, since $y$ anticommutes with $a$ and $b$ and thus commutes with $I = ab$, we have

$$\mathrm{rot}_\theta y = (\cos\frac{\theta}{2} - \sin\frac{\theta}{2} . I) y (\cos\frac{\theta}{2} + \sin\frac{\theta}{2} . I)$$

$$= y(\cos\frac{\theta}{2} - \sin\frac{\theta}{2} . I)(\cos\frac{\theta}{2} + \sin\frac{\theta}{2} . I) = y \; .$$

Now if $z$ is any vector in $E$ it can be decomposed into a component $z_{\parallel}$ in the plane of $I$ and a component $z_{\perp}$ orthogonal to that plane. Then we have

$$(2.5.3) \qquad \mathrm{rot}_\theta z = \mathrm{rot}_\theta z_{\parallel} + z_{\perp} = e^{-\frac{\theta}{2} I} z e^{\frac{\theta}{2} I} \; .$$

This formula gives the representation of a very simple isometry, namely that of a rotation in a two-dimensional Euclidean plane, by means of a bivector belonging to this plane. In chapter IV we will represent general isometries, both in the Euclidean and in the Lorentz case, in the form

$$Ax = e^{\frac{F}{2}} x \, e^{-\frac{F}{2}} ,$$

Where $F$ is a general Clifford bivector. Such a quantity $F = \sum_{i<k} F^{ik} e_i e_k$ has $\frac{1}{2}n(n-1)$ independent components, and thus exhibits the correct number of parameters.

# Chapter III

## Canonical representation of isometries.

3.1 We start this chapter by recalling some rather well known facts of a general character. Suppose we are given a finite dimensional vector space $E$ of dimension $n$ over a field $F$, which we shall assume to be either the field of the real numbers $R$ or the field of the complex numbers $K$. Small Greek letters will always denote elements of the field (but not conversely).

Let $A$ be a fixed linear transformation of $E$ into itself: $A E \subset E$, i.e. $A x \in E$ for all $x \in E$. Such a transformation is also called an endomorphism of $E$. A subspace $H$ of $E$ which is transformed into itself by $A$: $A H \subset H$, is called an invariant subspace of $E$, or, more precisely, an invariant subspace of $E$ relative to $A$. For a fixed vector $x \in E$, the subspace spanned by the vectors $x, Ax, A^2x, \ldots$ is clearly invariant.

A vector $x$ such that $Ax = \lambda x$ for some $\lambda \in F$ is called eigenvector of $A$ and $\lambda$ is called the eigenvalue corresponding to $x$. The one-dimensional subspace generated by an eigenvector $x$ is another example of an invariant subspace.

Given a polynomial $f(t) = \alpha_m t^m + \ldots + \alpha_0$ in an indeterminate $t$, we can form the linear transformation

$$f(A) = \alpha_m A^m + \ldots + \alpha_0 1,$$

where $1$ stands for the identity operator $1x = x$. For any vector $x \in E$ there exists a polynomial $g(t) \neq 0$ such that $g(A)x = 0$, because the $n + 1$ elements $x, Ax, A^2x, \ldots, A^nx$ are necessarily linearly dependent. This also shows that $g(t)$ is a polynomial of degree $m \leqslant n$ and that the leading coefficient

$\alpha_m$ can be chosen equal to 1. We say that $g(A)$ annihilates the vector $x$ if $g(A)x = 0$; for the sake of brevity we shall also say that the polynomial $g(t)$ annihilates $x$ if $g(A)$ does so. It is clear that the polynomials $g(t)$ which annihilate $x$ form an ideal in the ring of polynomials $F[t]$. It is well known that an ideal in this ring is always principal; hence, there exist polynomials $m_x(t)$ in the ideal which divide all $g(t)$ annihilating $x$. The $m_x(t)$ with leading coefficient 1 is called the minimal polynomial of $x$. One should keep in mind that $m_x(t)$ is defined relative to the linear transformation $A$.

Let $e_1, \dots, e_n$ be a basis of $E$ and $g_1(t), \dots, g_n(t)$ corresponding annihilating polynomials. Clearly the product $G(t) = g_1(t) \dots g_n(t)$ annihilates the whole space $E$, that is $G(A)x = 0$ for all $x \in E$. Hence the polynomials which annihilate $E$ form a non zero ideal in $F[t]$ and their greatest common divisor $m(t)$ with leading coefficient 1 (the generator of the ideal), will be called the minimal polynomial of the space $E$ (relative to $A$). We show that $m(t)$ is equal to the least common multiple (l.c.m.) $\bar{m}(t)$ of the minimal polynomials of the basis elements $e_i$. Indeed $\bar{m}(t)$ annihilates $E$ and hence $m(t)$ divides $\bar{m}(t)$. On the other hand $m(t)$ annihilates every $e_i$, hence every $m_{e_i}(t)$ divides $m(t)$ and $\bar{m}(t)$ divides $m(t)$.

The minimal polynomial $m(t)$ of $E$ has degree $\leq n$. This is an immediate consequence of the fact, to be proved now, that there exists a vector $x \in E$ whose minimal polynomial is $m(t)$.

Lemma. Let $x$ and $y$ be two non collinear vectors in $E$ and let $m_x(t)$ and $m_y(t)$ be their minimal polynomials. Let $\bar{m}(t)$ be the least common multiple of $m_x(t)$ and $m_y(t)$. Then there exists a vector $z$ in the two-dimensional space spanned by $x$ and $y$ whose minimal polynomial $m_z(t)$ is $\bar{m}(t)$.

Let us observe that this lemma is exactly the above statement in the case $n = 2$.

__Proof.__ Every vector $u = \xi x + \eta y$ has a minimal polynomial $m_u(t)$ which divides $\bar{m}(t)$, since $\bar{m}(t)$ annihilates both $x$ and $y$ and hence also $u$. Now $\bar{m}(t)$ has only a finite number of divisors, hence there must exist two non collinear vectors $u_1$ and $u_2$ whose minimal polynomials are equal, let us call it $m_u(t)$. We can write $x = \alpha_1 u_1 + \beta_1 u_2$, $y = \alpha_2 u_1 + \beta_2 u_2$, and we have $m_u(A)x = 0$, $m_u(A)y = 0$. Hence $m_x(t)$ and $m_y(t)$ both divide $m_u(t)$ and thus $\bar{m}(t)$ divides $m_u(t)$. But on the other hand, $m_u(t)$ divides $\bar{m}(t)$ and since the leading coefficient of both polynomials is 1, we get $m_u(t) = \bar{m}(t)$. Q.E.D.

It now follows by induction that if $e_1, \ldots, e_n$ are $n$ linearly independent vectors in $E$, then there exists a vector $z$ in $E$ whose minimal polynomial is the l.c.m. of the $m_{e_i}(t)$, i.e. $m_z(t) = m(t)$.

3.2 Suppose now that $m(t)$ is written as a product

$$m(t) = \prod_{i=1}^{s} m_i(t),$$

where the polynomials $m_i(t)$ are __relatively prime in pairs.__ In the applications the $m_i(t)$ will be taken as the powers of distinct prime polynomials.

We propose to __decompose__ the space $E$ into a direct sum of invariant subspaces $E_i$ such that for each $i$ the __minimal polynomial of__ $E_i$ __is__ $m_i(t)$.

Let us write

$$M_k(t) = \frac{m(t)}{m_k(t)} = \prod_{i \neq k} m_i(t).$$

We assert that the polynomials $M_k$ have g.c.d. 1. We prove this by induction. For $s = 2$ the statement is obviously true. Put now

$$M'_k(t) = \prod_{i \neq k,s} m_i(t) \qquad (k = 1,2,\ldots, s-1).$$

Then by hypothesis the g.c.d. of the $M'_k(t)$ is 1; hence the g.c.d. of the $M_k(t) = M'_k(t)\, m_s(t)$ $(k = 1,\ldots, s-1)$ is $m_s(t)$. But $m_s(t)$ is relatively prime to $M_s(t)$, hence the g.c.d. of $M_1(t),\ldots, M_s(t)$ is 1. It follows now that there exist polynomials $g_i(t)$ $(i = 1,\ldots, s)$ such that

$$\text{(3.2.1)} \qquad \sum_{i=1}^{s} g_i(t)\, M_i(t) = 1,$$

and consequently

$$\text{(3.2.2)} \qquad \sum_{i=1}^{s} g_i(A)\, M_i(A) = 1.$$

We now define $E_i$ $(i = 1,\ldots, s)$ to be the set of those vectors of $E$ which are annihilated by $m_i(t)$. It is obvious that $E_i$ is an invariant subspace of $E$ and we shall show later that $m_i(t)$ is actually the minimal polynomial of $E_i$. We first prove that $E_i = g_i(A)\, M_i(A)\, E$. Indeed the right hand side is annihilated by $m_i(A)$, since $m_i(A)\, g_i(A)\, M_i(A)\, E = g_i(A)\, m(A)\, E = 0$, thus $g_i(A)\, M_i(A)\, E \subset E_i$. On the other hand, suppose that $y \in E_i$, that is $m_i(A)\, y = 0$. Then $M_k(A)\, y = 0$ for all $k \neq i$, since $m_i(t)$ divides $M_k(t)$. From this and (3.2.2) it follows that

$$y = \left( \sum_{k=1}^{s} g_k(A) M_k(A) \right) y = g_i(A) M_i(A) y \in g_i(A) M_i(A) E,$$

and thus that $E_i \subset g_i(A) M_i(A) E$.

In the same way it follows from (3.2.2) that every $x \in E$ can be decomposed into the sum of $x_i \in E_i$, $i = 1, \ldots, s$. Indeed, we can write

$$(3.2.3) \qquad x = \sum_{i=1}^{s} x_i, \quad \text{with} \quad x_i = g_i(A) M_i(A) x \in E_i.$$

The last formula gives a decomposition of the space $E$ into the sum of the subspaces $E_i$. Let us prove that $E$ is the direct sum of the $E_i$. To this end we have to show that if

$$(3.2.4) \qquad \sum_{k=1}^{s} y_k = 0, \quad y_k \in E_k,$$

then $y_k = 0$ for all $k$. As shown above, $y_i \in E_i$ implies $y_i = g_i(A) M_i(A) y_i$. On the other hand (3.2.4) gives $y_i = -\sum_{k \neq i} y_k$. Now $M_i(A) y_k = 0$, if $k \neq i$, since $m_k(A) y_k = 0$ and $m_k(t)$ divides $M_i(t)$. Hence $y_i = 0$.

Now we shall prove that the minimal polynomial of $E_i$ is $m_i(t)$. We already know that $m_i(t)$ annihilates $E_i$. If $E_i$ were annihilated by a proper divisor $h(t)$ of $m_i(t)$, then $E$ would be annihilated by $h(t) M_i(t)$, which is a proper divisor of $m(t)$.

The remainder of this section may be omitted at a first reading. As a matter of fact, it is not needed in the sequel except for an incidental application in section 4.8.

All congruences here are taken modulo $m(t)$. Set

$$(3.2.5) \qquad e_i(t) \equiv g_i(t) M_i(t),$$

then

(3.2.6) $$\sum_k e_k(t) \equiv 1.$$

Since $m(t) \mid M_i(t)\ M_k(t)$ when $i \neq k$, we have

(3.2.7) $$e_i(t)\, e_k(t) \equiv 0 \text{ when } i \neq k,$$

and hence

$$e_i(t) = e_i(t)\,.\,1 \equiv e_i^2(t),$$

that is, $e_i(t)$ is idempotent. Furthermore

(3.2.8) $$e_i(A) = \begin{cases} 1 & \text{in } E_i \\ 0 & \text{in } E_k,\ k \neq i, \end{cases}$$

and $\sum e_k(A) = 1$ in $E$.

Let $f(t)$ be a polynomial, and set

(3.2.9) $$f_i(t) \equiv f(t)\, e_i(t).$$

Then

(3.2.10) $$f_i(A) = \begin{cases} f(A) & \text{in } E_i \\ 0 & \text{in } E_k,\ k \neq i \end{cases} \qquad \text{and } f(A) = \sum f_k(A) \text{ in } E.$$

Formulae (3.2.9) and (3.2.10) give a kind of spectral decomposition of the identity and of $f(A)$.

More generally, let $h^k(t)$ be any polynomials, and set

(3.2.11) $$h(t) = \sum_k h^k(t)\, e_k(t).$$

Then we have the "interpolation formula"

(3.2.12) $$h(A) = h^i(A) \text{ in } E_i,\quad i = 1, 2, \ldots, s.$$

The polynomial $h(t)$ is uniquely determined by the additional assumption that its degree is less than that of $m(t)$.

The inverse $A^{-1}$ exists if there is no non zero subspace annihilated by A, that is if $m(o) \neq 0$. Introduce in this case the polynomial

$$\bar{m}(t) = \frac{m(t) - m(o)}{m(o).t} .$$

Then $m(A) = 0$ can be written $A\,\bar{m}(A) + 1 = 0$, hence $A^{-1} = -\bar{m}(A)$, a polynomial in A, and the above scheme can be applied. For an isometry, as defined in the next section, the inverse always exists.

This scheme can also be applied if $f(t)$ in (3.2.9) or $h^k(t)$ in (3.2.11) are Laurent polynomials which may contain both positive and negative powers of t.

3.3 We now provide our n-dimensional vector space E with a metric, i.e., we postulate the existence of symmetric bilinear form $(x \mid y)$, defined for every couple $(x,y)$ of elements of E and taking values in the field F, and having the properties

$$(x \mid y) = (y \mid x),$$

$$(\lambda x_1 + \mu x_2 \quad y) = \lambda (x_1 \mid y) + \mu (x_2 \mid y).$$

Two vectors x and y are said to be orthogonal if $(x \mid y) = 0$. A non zero vector satisfying $(x \mid x) = 0$, that is a vector orthogonal to itself, is called isotropic. A metric is called regular if for every $x \neq 0$ there exists a y such that $(x \mid y) \neq 0$, i.e. if for every vector x there exists a vector y which is not orthogonal to it. A metric is called singular if it is not regular, i.e. if there exists a vector $y \neq 0$ such that $(x \mid y) = 0$ for all $x \in E$ i.e. if there exists $y \in E$ orthogonal to E.

If $e_1, \ldots, e_n$ is a basis of E, and we write

$$x = \sum_{i=1}^{n} x_i e_i , \qquad y = \sum_{i=1}^{n} y_i e_i ,$$

then we get

(3.3.1) $$(x \mid y) = \sum_{i,j=1}^{n} x_i \, y_j \, (e_i \mid e_j).$$

If we set $g_{ij} = (e_i \mid e_j)$, the metric can be written

(3.3.2) $$(x \mid y) = \sum_{i,j=1}^{n} g_{ij} \, x_i \, y_j.$$

We have $g_{ij} = g_{ji}$ and it is easy to see that $(x \mid y)$ is regular if and only if $\det (g_{ij}) \neq 0$.

Our two main particular cases will be those of the Euclidean metric and the Lorentz metric. These, with suitable bases, can be written in the following standard forms.

a) the Euclidean metric:

(3.3.3) $$(x \mid x) = x_1^2 + x_2^2 + \dots + x_n^2,$$

that is $g_{ii} = 1$ and $g_{ij} = 0$ if $i \neq j$ ;

b) the Lorentz metric:

(3.3.4) $$(x \mid x) = x_1^2 - x_2^2 - \dots - x_n^2,$$

that is $g_{11} = 1$, $g_{ii} = -1$ if $i = 2, \dots, n$, $g_{ij} = 0$ if $i \neq j$.

Both metrics are obviously regular.

A subspace H of E (which can be E itself) will be called regular if the restriction of $(x \mid y)$ to H is regular, i.e. if for every $x \in H$ there exists a $y \in H$ with $(x \mid y) \neq 0$. A subspace H of E is called singular if there exists $y \in H$, $y \neq 0$, orthogonal to H. Such a y is isotropic, since it is orthogonal to itself. A non zero subspace H is called totally isotropic if $(x \mid x) = 0$ for all $x \in H$ i.e. all its vectors are isotropic. This is equivalent to $(x \mid y) = 0$ for all $x, y \in H$ by the relation

(3.3.5) $$2(x \mid y) = (x + y \mid x + y) - (x \mid x) - (y \mid y).$$

Obviously totally isotropic implies singular but not vice versa. But a singular subspace always contains at least one isotropic vector, and hence the totally isotropic subspace spanned by this vector. We observe that any subspace of a totally isotropic subspace is totally isotropic, whereas a subspace of a singular subspace can be regular.

Examples of isotropic vectors and singular subspaces in the real case can be found on p. 109 ; examples of totally isotropic subspaces of dimension $>1$ over the complex field can be found on p. 104.

The set of vectors of a subspace H of E which are orthogonal to H forms a subspace of H called the radical of H, which is denoted by rad H. This subspace is either 0 or totally isotropic since it is orthogonal to itself. Obviously H is regular if and only if rad $H = 0$ and H is totally isotropic if and only if $H = \text{rad } H \neq 0$.

A subspace H which is not totally isotropic can be decomposed into the direct sum of rad H and some subspace G of H in infinitely many ways. Every such G is regular. Indeed, if $v \in G \subset H$ and $v \perp G$, then $v \perp H = G + \text{rad } H$, since $v \perp \text{rad } H$. Thus $v \in G \cap \text{rad } H$, i.e. $v = 0$.

This result can be interpreted in terms of the quotient space H/ rad H. In the first place the scalar product $(x \mid y)$ defines a metric on H/ rad H, since $x \equiv \tilde{x}$ (mod. rad H), $y \equiv \tilde{y}$ (mod. rad H) imply $(x \mid y) = (\tilde{x} \mid \tilde{y})$. The above result expresses the fact that the quotient space H/rad H is regular with respect to this metric.

Let H be a subspace of a (regular or singular) vector space E. We define the orthogonal complement $H^{\perp}$ of H in E as the subspace of E consisting of all vectors in E which are orthogonal to H, that is orthogonal to all vectors in H. By means of this concept the

the radical of H can be expressed

(3.3.6) $\quad \operatorname{rad} H = H \cap H^{\perp}$.

In order to investigate the properties of the orthogonal complement we first recall the well known fact that a system of linear homogeneous equations in which the number of the equations is less than the number of the unknowns always has non trivial solutions. (This fact is obvious for one equation, and the general case then follows by elimination and induction.)

One subsequent results will be the consequences of the following theorem.

*If* $H$ *is a proper subspace of* $E$, *then* $H^{\perp} \neq 0$. *Proof*: Let $e_1, e_2, \ldots, e_n$ be a basis of $E$ such that $e_1, \ldots, e_m$, with $m < n$, is a basis for $H$. Set $u = \lambda_1 e_1 + \ldots + \lambda_n e_n$ and solve the $m$ equations $(u \mid e_k) = 0$, $k = 1, \ldots, m$, for the $n$ unknowns $\lambda_1, \ldots, \lambda_n$.

Now we prove the following. *If* $H$ *is a proper subspace of* $E$ *and* $H$ *is regular, then* $E$ *is the direct sum of* $H$ *and* $H^{\perp}$.

*Proof*. If $x \in H$, we have the decomposition $x = x + 0$, $x \in H$, $0 \in H^{\perp}$. If $x \in E$, but $x \notin H$, denote by $\tilde{E}$ the space spanned by $x$ and $H$. Then $H$ is a proper subspace of $\tilde{E}$, thus by virtue of the preceding theorem, there exists $v \in \tilde{E}$, $v \perp H$, $v \neq 0$. Since $H$ is regular, $v \notin H$, and thus $v = \lambda x - z$, with $\lambda \neq 0$ and $z \in H$. By setting $\lambda^{-1} v = u$, $\lambda^{-1} z = y$ we obtain $x = y + u$ with $y \in H$, $u \in H^{\perp}$. The proof of $E = H + H^{\perp}$ herewith is complete. Since $H$ is regular, $H \cap H^{\perp} = 0$ and hence the sum is direct. The vector $y$ is called the *orthogonal projection* of the vector $x \in E$ into the subspace $H$.

As a consequence: if H is regular, $\dim E = \dim H + \dim H^{\perp}$.

In our further considerations about the orthogonal complement of a subspace of E with respect to E, we suppose that E is regular.

We first show that if $\{e_k\}$ is a basis of a regular space E of dimension n, then there exists a basis $\{e^k\}$ of E, satisfying the equations

$$(e_i \mid e^k) = \delta_i^k = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases} \tag{3.3.7}$$

The two bases are said to be complementary to each other and they determine each other in a unique way. To find such a basis we let $E_k$ be the space spanned by the vectors $e_1, \ldots, e_n$ with the exception of the vector $e_k$. Then $E_k$ is a proper subspace of E, thus there exists in E a non zero vector $v \perp E_k$. By virtue of the regularity of E, we have $(v \mid e_k) = \mu \neq 0$, and $e^k = \mu^{-1} v$ is a solution of the equations (3.3.7) for $i = 1,2,\ldots, n$. The uniqueness of the solutions $e^k$ follows from the regularity of the metric, and their linear independence is obvious from the equations themselves. Hence the $e^k$ form a basis of E. A glance at the equations shows that the relation between the bases $\{e_k\}$ and $\{e^k\}$ is reciprocal.

For the sake of completeness we give some explicit formulae. As on p. 91, we set $g_{ik} = g_{ki} = (e_i \mid e_k)$. Since E is regular with respect to the metric, $\det(g_{ik}) \neq 0$ and the matrix $(g^{ik}) = (g_{ik})^{-1}$ exists. We have

$$e^k = \sum_{\tau=1}^{n} g^{k\tau} e_\tau, \quad e_k = \sum_{\tau=1}^{n} g_{k\tau} e^\tau. \tag{3.3.8}$$

We return to the orthogonal complement of a subspace H of E and prove:

If E is regular, then $H^{\perp\perp} = H$ and $\dim H + \dim H^{\perp} =$

$=$ dim E.

Proof. The statement is obvious if $H = 0$ or $H = E$. Next let H be of dimension m, where $o < m < n =$ the dimension of E. The basis $e_1, \ldots, e_n$ considered above may be chosen so that $e_1, \ldots, e_m$ is a basis of H. In this case $e^{m+1}, \ldots, e^n$ is a basis $H^{\perp}$. Hence $H^{\perp\perp} = H$ and the assertion about the dimensions immediately follows.

As a consequence of this result we get rad $H^{\perp} = H^{\perp} \cap H^{\perp\perp} = H^{\perp} \cap H =$ rad H; this implies that H and $H^{\perp}$ are regular or singular simultaneously.

When $H^{\perp\perp} = H$, we say that H and $H^{\perp}$ are the orthogonal complements of one another.

The last theorem dealing with the concept of the orthogonal complement will be the following.

Let E be regular and H a subspace of E. Then $H \cap H^{\perp}$ and $H + H^{\perp}$ are the orthogonal complements of one another.

Proof. $H \cap H^{\perp}$ is orthogonal to both H and $H^{\perp}$, since it is contained in both $H^{\perp}$ and H. Consequently it is also orthogonal to $(H + H^{\perp})$, and therefore $(H \cap H^{\perp}) \subset (H + H^{\perp})^{\perp}$. On the other hand the latter subspace is orthogonal to $H + H^{\perp}$ and hence to H and $H^{\perp}$. Therefore $(H + H^{\perp})^{\perp}$ is contained in $H^{\perp}$ and in H, and hence in $(H \cap H^{\perp})$; that is $(H + H^{\perp})^{\perp} \subset (H \cap H^{\perp})$.

In the case that H is regular, rad $H = 0$ and the above result is in agreement with an older one, namely, that, in this case, $H + H^{\perp} = E$.

A linear transformation L of a vector space E, provided with a metric $(x \mid y)$, into itself is called an automorphism of the metric or an isometry with respect to this metric, if

$$(3.3.9) \qquad (Lx \mid Ly) = (x \mid y)$$

for all $x, y \in E$. It is enough to suppose that

$$(Lx \mid Lx) = (x \mid x)$$

for all $x \in E$ by virtue of the relation (3.3.5). If the metric is regular, then L is a one-to-one transformation of E onto itself. In fact if $Ly = 0$, then it follows from (3.3.9) that $(x \mid y) = 0$ for all $x \in E$, and hence that $y = 0$. Thus $\dim L(E) = n$ and L is onto. It follows that $L^{-1}$ exists, that is $LL^{-1} = L^{-1}L = 1$ and $L^k$ is an isometry for all (positive or negative) integers k.

Let H be an invariant subspace with respect to the isometry L of the regular metric $(x \mid y)$. Then L maps H onto H since $\dim LH = \dim H$ and $LH \subset H$. $H^{\perp}$ is also invariant with respect to L. Indeed, if $y \in H^{\perp}$ and $x \in H$, then $x = L\tilde{x}$, for some $\tilde{x} \in H$ and

$$(Ly \mid x) = (Ly \mid L\tilde{x}) = (y \mid \tilde{x}) = 0,$$

thus $Ly \in H^{\perp}$. We also note that if $f(t)$ is any polynomial then H is invariant with respect to $f(L)$ or to $f(L^{-1})$ (or to a Laurent-polynomial).

A regular invariant subspace of E is called reducible if it contains an invariant regular proper subspace, and irreducible otherwise.

A reducible subspace can be written as the direct sum of two invariant regular proper subspaces, which are orthogonal to each other, by what has been said above. Because of the finiteness of the dimension, a regular invariant subspace is either irreducible or contains an irreducible proper subspace.

A fundamental problem consists of decomposing a space E with a given metric and a given isometry L of this metric into subspaces invariant and irreducible with respect to L. We will carry out such decompositions in two important cases: the real Euclidean space and

and the real Lorentz space.

To this end we shall prove in the next section a few lemmas on general metrics, which present some independent interest in themselves.

3.4 Let E be an n-dimensional vector space and let A be an endomorphism of E. We first collect a few properties of invariant subspaces, which will play an important role in this section. Let f (t), g (t), h (t) be arbitrary polynomials. A subspace of E invariant under A is invariant under f(A). The subspace annihilated by g(A) is invariant under A. The subspace h(A) E is invariant under A. The degree of the minimal polynomial with respect to A of a subspace H invariant under A is $\leq$ dim H.

To make things clearer, it is sometimes useful to talk of the restriction of A to an invariant subspace H meaning A as it acts on the elements of H only.

Let now E have a metric $(x \mid y)$ and consider an isometry L of this metric. In Lemmas I - VII we do not require that $(x \mid y)$ be regular, but we do suppose that L possesses an inverse $L^{-1}$.

Lemma I - If $Lx = \lambda x$, $Ly = \mu y$ and $\lambda\mu \neq 1$, then $(x \mid y) = 0$.

Proof. $0 = (Lx \mid Ly) - (x \mid y) = (\lambda\mu - 1)\ (x \mid y)$.

Lemma II - For any polynomial f(t), we have

$$(f(L)\, x \mid y) = (x \mid f(L^{-1})y).$$

Proof. We have $(L^k x \mid L^k z) = (x \mid z)$ for all integers k. Setting $z = L^{-k} y$ we obtain $(L^k x \mid y) = (x \mid L^{-k} y)$. From this our assertion follows.

Lemma III - If f(t) and g(t) are relatively prime polynomials, $f(L)\, x = 0$, $g(L^{-1})\, y = 0$, then $(x \mid y) = 0$.

Proof. There exist two polynomials $a(t)$ and $b(t)$ such that

$$a(t)\, f(t) + b(t)\, g(t) = 1.$$

Then

$$x = \left[a(L)\, f(L) + b(L)\, g(L)\right] x = b(L)\, g(L)\, x,$$

since $f(L)\, x = 0$. Hence by Lemma II

$$(x \mid y) = (b(L)\, g(L)\, x \mid y) = (b(L)\, x \mid g(L^{-1})y) = 0, \text{ since } g\, (L^{-1})y = 0.$$

Example - Lemma I follows from Lemma III by setting $f(t) = t - \lambda$, $g(t) = t - \mu^{-1}$.

Let us recall the definition of reciprocal polynomials. The polynomials $f(t) = \sigma \sum_{k=0}^{\tau} \alpha_k t^k$ and $\tilde{f}(t) = \tau \sum_{k=0}^{\tau} \alpha_{\tau-k} t^k$ with $\alpha_0 \alpha_\tau \neq 0, \sigma\tau \neq 0$ are said to be reciprocal to each other. Obviously $\sigma \tilde{f}(t) = \tau\, t^{\tau}\, f(t^{-1})$ and $\tau f(t) = \sigma\, t^{\tau} \tilde{f}\, (t^{-1})$.

With this terminology we can reformulate Lemma III:

Lemma III$^{\text{bis}}$. - If $f(t)$ and $\tilde{g}(t)$ are relatively prime, the subspaces of E annihilated by $f(L)$ and $g(L)$ respectively are orthogonal to each other.

As a corollary we have:

Lemma IV - If the mutually reciprocal polynomials $f(t)$ and $\tilde{f}(t)$ are relatively prime, the subspace of E annihilated by $f(L)$ ( or by $\tilde{f}(L)$ ) is either 0 or totally isotropic.

Example I - Let $f(t) = (t - \lambda)^{\tau}$ with $\lambda^2 \neq 1$, i.e. $\lambda \neq \frac{1}{\lambda}$. Hence $(t - \lambda)^{\tau}$ and $t^{\tau}(t^{-1} - \lambda)^{\tau} = (-\lambda)^{\tau}\ (t - \lambda^{-1})^{\tau}$ are relatively prime. Consequently the space annihilated by $(L - \lambda)^{\tau}$ with $\lambda^2 \neq 1$ is totally isotropic or 0.

Example II - Let $Q\,(t) = t^2 + \beta t + \gamma$ be a positive quadratic polynomial with real coefficients. If $\gamma \neq 1$, $Q(t)$ and $\tilde{Q}(t)$ are relatively prime and so are $f(t) = \left[Q(t)\right]^{\tau}$ and $\tilde{f}(t) =$

$= [\tilde{Q}(t)]^{\dagger}$ . Hence if $\gamma \neq 1$, the subspace annihilated by $f(L)$ is totally isotropic or 0.

Lemma V - If $f(t)$ is self-reciprocal, the relations $f(L)\, x = 0$ and $f(L^{-1})\, x = 0$ imply each other.

Proof. Clear.

Example I - $f(t) = (t - \varepsilon)^{p}$ with $\varepsilon^{2} = 1$.

Example II - $f(t) = [Q(t)]^{p}$ with $Q(t) = t^{2} + \beta t + 1$.

Lemma VI - Let $g(t)$ be self-reciprocal and $[g(t)]^{p}$, $p \geqslant 2$, be the minimal polynomial of a certain invariant subspace H of E. Let $q = \left[\frac{p+1}{2}\right]$ ( $[\xi]$ is the largest integer $\leqslant \xi$ ). Then $H' = [g(L)]^{q} H$ is totally isotropic.

Proof. Since $q < p$ it follows that $H' \neq 0$. Let $x \in H'$. Then $x = [g(L)]^{q} u$, $u \in H$ and according to Lemma II $(x \mid x) = (w \mid u)$, where $w = [g(L^{-1})]^{q} [g(L)]^{q} u$. But $[g(L)]^{q} [g(L)]^{q} u =$ $= [g(L)]^{2q} u = 0$, since $2q \geqslant 0$ and so $w = o$ by Lemma V.

Example I - $g(t) = t - \varepsilon$ with $\varepsilon^{2} = 1$. The totally isotropic subspace $H' = (L - \varepsilon)^{q} H$ has the minimal polynomial $(t - \varepsilon)^{p-q}$. Consequently $\dim H' \geqslant p - q$. Now $q = \frac{p+1}{2}$ if $p$ is odd and $= \frac{p}{2}$, if $p$ is even. Consequently $\dim H' \geqslant \frac{p-1}{2}$, if $p$ is odd and $\geqslant \frac{p}{2}$, if $p$ is even.

Example II - $g(t) = t^{2} + \beta t + 1 \geqslant 0$, $\beta$ real.

Lemma VII - Let $\varepsilon^{2} = 1$. Suppose that some invariant subspace H of E has the minimal polynomial $(t - \varepsilon)^{p}$ with $p > 1$. Then the subspace $E'$ of E annihilated by $L - \varepsilon$ is singular.

Proof. We shall make use of the fact, implied by the hypothesis, that there exists in E a vector $x$ whose minimal polynomial is $(t - \varepsilon)^{p}$ with $p > 1$. Let $z = (L - \varepsilon)^{p-1} x \neq 0$

and let y be *arbitrary* in $E'$. Then, according to Lemma II, $(z \mid y) = (x \mid w)$, with $w = (L^{-1} - \varepsilon)^{\rho-1} y$. Now $(L - \varepsilon)^{\rho-1} y = 0$, since $(L - \varepsilon)y = 0$ and $\rho - 1 \geqslant 1$. Thus, in view of Lemma V, $w = 0$; and hence $(z \mid y) = 0$. On the other hand $(L - \varepsilon)z = (L - \varepsilon)^{\rho} x = 0$, which means that $z \in E'$ and so $E'$ is singular.

*Lemma VIII* - Let E be *regular* with respect to the metric $(x \mid y)$. Let $f(t)$ be a polynomial, $G = f(L)\,E$ and H be the subspace annihilated by $f(L^{-1})$. Then G and H are the orthogonal complements of one another.

*Proof*. Let u be arbitrary in G and v fixed in H. Then $u = f(L)\,x$, with $x \in E$, and $f(L^{-1})\,v = 0$. According to Lemma II, $(u \mid v) = (f(L)\,x \mid v) = (x \mid f(L^{-1})\,v) = 0$. Hence $v \in G^{\perp}$, that is $H \subset G^{\perp}$. Conversely, let $(u \mid w) = 0$ for a fixed $w \in G^{\perp}$ and an arbitrary $u \in G$. This means $(f(L)\,x \mid w) = 0$ and hence $(x \mid f(L^{-1})w) = 0$ for all $x \in E$. By the regularity of E it follows that $f(L^{-1})w = 0$, that is $w \in H$, $G^{\perp} \subset H$. Consequently $G^{\perp} = H$. Since $G^{\perp\perp} = G$, it follows that $H^{\perp} = G$.

Using Lemma V. we have as a corollary:

*Lemma VIII$^{\text{bis}}$* - Let E be regular, $f(t)$ self-reciprocal, G the subspace $f(L)\,E$ and H the subspace which is annihilated by $f(L)$. Then G and H are the orthogonal complememts of each other.

*Example*. - $f(t) = (t - \varepsilon)^{\rho}$.

3.5 We shall now focus our attention on the most familiar particular case, that of a *Euclidean metric* over a real vector space E of dimension n, introduced in formula (3.3.3.). As is well known, such a metric can also be defined by *the intrinsic property that* $(x \mid x) \geqslant 0$ *and* $(x \mid x) = 0$ only for $x = 0$. This implies that *every subspace of a Euclidean space is regular*.

We show that under the above assumptions a basis $e_1, \dots, e_n$ of E can be chosen so that for $x = x_1 e_1 + \dots + x_n e_n$ we have

$$(x \mid x) = x_1^2 + x_2^2 + \dots + x_n^2 .$$

The proof proceeds by induction on the number of dimensions. If $n = 1$, let $0 \neq a \in E$, $(a \mid a) = \alpha^2 > 0$, hence taking $e_1 = \alpha^{-1} a$, we obtain a unit vector, i.e. $(e_1 \mid e_1) = 1$. Then for any $x = x_1 e_1 \in E$ we have $(x \mid x) = x_1^2$. If E has dimension $n > 1$, choose a one dimensional subspace H. Its complement $H^{\perp}$ has dimension $n - 1$. Writing $x = x' + x''$, $x' \in H$, $x'' \in H^{\perp}$, we obtain $(x \mid x) = (x' \mid x') + {}$ ${} + (x'' \mid x'')$, from which our assertion follows by induction.

Let now L be an isometry of the Euclidean metric. We propose to seek the the "canonical decomposition" of L. The minimal polynomial of E relative to L is the product of distinct factors of the form $(t - \lambda)^{\rho}$ and $[Q(t)]^{\tau}$, where $Q(t) = t + \beta t + \gamma$ ($\lambda, \beta, \gamma$ real) is irreducible over the real field. We assert that the only possible values are

$$\lambda = \pm 1, \quad \gamma = 1, \quad \rho = 0 \text{ or } 1, \quad \tau = 0 \text{ or } 1.$$

(The conditions $\rho \leq 1$, $\tau \leq 1$, mean, in learned terms, that the elementary divisors are all simple.) Indeed, since E has no totally isotropic subspace (because there are no isotropic vectors!), $\lambda = \pm 1$ and $\gamma = 1$ follow from Examples I. and II. for Lemma IV in 3.4 and $p = 1, \tau = 1$ follow from Examples I. and II. for Lemma VI. in 3.4. Furthermore $|\beta| < 2$, because $Q(t)$ is irreducible.

By the results of 3.2, the space E is the direct sum of the subspaces, whose minimal polynomials are the actually ocurring factors $t - 1$, $t + 1$ and the various $Q(t) = t^2 + \beta t + 1$ of the minimal polynomial of E. We shall denote these subspaces by $E_1$, $E_{-1}$ and

$E_Q$, whenever they occur in the decomposition of E. It follows from Lemma $III^{bis}$ of 3.4 that these subspaces are mutually orthogonal.

In the sequel, whenever we discuss a subspace, we suppose that it actually occurs in the decomposition of E.

All vectors of $E_1$ are left invariant by L; all vectors of $E_{-1}$ undergo a symmetry in the origin. According as $E_{-1}$ is of even or odd dimension, L is called a rotation or a reflection. If any of these two subspaces has dimension $\geqslant 2$, it can be decomposed in infinitely many ways into the direct sum of mutually orthogonal invariant lines.

Consider a subspace $E_Q$ and let $x \in E_Q$ be an arbitrary vector. Setting $\beta = -2\cos\theta$, with $|\cos\theta| < 1$, we have $(L^2 - 2\cos\theta\, L + 1)x = 0$, or

(3.5.1) $$(L + L^{-1})\, x = 2\cos\theta \,.\, x.$$

Write (3.5.1) for a particular non zero vector $a \in E_Q$, that is

(3.5.2) $$(L + L^{-1})\, a = 2\cos\theta \,.\, a,$$

and define the vector b by

(3.5.3) $$(L - L^{-1})\, a = 2\sin\theta \,.\, b,$$

where $\sin\theta \neq 0$ is determined except for sign by the equation $\cos\theta = -\frac{1}{2}\beta$. Unless otherwise stated we take $\sin\theta$ positive.

Since $(a \mid La) = (a \mid L^{-1} a)$ by Lemma II, we have

(3.5.4) $$(a \mid b) = 0.$$

On the other hand, from (3.5.2) and (3.5.3),

$$\cos^2\theta \,.\, (a \mid a) + \sin^2\theta \,.\, (b \mid b)$$
$$= \frac{1}{4}\left( 2(La \mid La) + 2(L^{-1} a \mid L^{-1} a) \right) = (a \mid a),$$

and hence

(3.5.5) $$(a \mid a) = (b \mid b).$$

Setting $x = b$ in (3.5.1), we obtain

(3.5.6) $$(L + L^{-1})\, b = 2\cos\theta \,.\, b.$$

Applying to (3.5.2) and (3.5.3) the respective operators $(L + L^{-1})$ and $(L - L^{-1})$ and subtracting we find

$$4a = 2\cos\theta . (L + L^{-1})a - 2\sin\theta . (L - L^{-1})\, b,$$

which by means of (3.5.2) reduces to

$$(L - L^{-1})\, b = -2\sin\theta . a. \qquad (3.5.7)$$

Combining formulae (3.5.2) - (3.5.7), we obtain

$$(3.5.8)\begin{cases} La = \cos\theta . a + \sin\theta . b & L^{-1} a = \cos\theta . a - \sin\theta . b \\ Lb = -\sin\theta . a + \cos\theta . b & L^{-1} b = \sin\theta . a + \cos\theta . b \\ (a \mid a) = (b \mid b), & (a \mid b) = 0. \end{cases}$$

If $E_Q$ is a two-plane, the formulae (3.5.8) describe a rotation in that plane. If $\dim E_Q > 2$, $E_Q$ can, in infinitely many ways, be decomposed into the direct sum of mutually orthogonal invariant two-planes undergoing the same rotation as the one exhibited by the formulae (3.5.8). As a by-product we find the fact that the dimension of $E_Q$ is even. For the proof we refer to 4.2, where the same question is treated in a slightly different language. All the invariant two-planes have the minimal polynomial $Q(t)$.

After this somewhat sophisticated treatment in the real, let us see the complex treatment. We pass from the real subspace $E_Q$ to the complex subspace $E_Q + i E_Q$, whose elements are the vectors $x + iy$ with $x \in E_Q$, $y \in E_Q$. Complex numbers are admitted as scalars and multiplication is defined by

$$(\lambda + i\mu)(x + iy) = \lambda x - \mu y + i(\mu x + \lambda y).$$

$E_Q$ can be considered as a subset of $E_Q + i E_Q$.

The minimal polynomial $Q(t)$, which was irreducible over R, now decomposes into two linear factors

$$Q(t) = t^2 - 2\cos\theta . t + 1 = (t - e^{-i\theta})(t - e^{i\theta}).$$

According to the general theory of section 3.2, $E_Q + i\,E_Q$ splits into the direct sum of two subspaces $H$ and $\overline{H}$, which are annihilated by $t - e^{-i\theta}$ and $t - e^{i\theta}$, respectively.

It is clear that if $z = a + i\,b \in H$, then $\overline{z} = a - ib \in \overline{H}$ and vice versa. We then have

$$(3.5.9) \qquad Lz = e^{-i\theta} z, \quad L\overline{z} = e^{i\theta}\,\overline{z}.$$

The vectors $z$ and $\overline{z}$ are isotropic since

$$(z|z) = (Lz\,|\,Lz) = (e^{-i\theta}z\,|\,e^{-i\theta}z) = e^{-2i\theta}\,(z\,|\,z)$$

and $e^{-2i\theta} \neq 1$ implies $(z\,|\,z) = 0$; similarly $(\overline{z}\,|\,\overline{z}) = 0$. We have

$$(a + ib\,|\,a + ib) = (a\,|\,a) - (b\,|\,b) - 2i\,(a\,|\,b) = 0,$$

and hence $(a\,|\,a) = (b\,|\,b)$, $(a\,|\,b) = 0$. Furthermore

$$L\,(a + ib) = e^{-i\theta}(a + ib) = (\cos\theta - i\sin\theta)\,(a + ib)$$

$$= \cos\theta . a + \sin\theta . b + i\,(\cos\theta . b - \sin\theta . a),$$

and so again

$$L\,a = \cos\theta . a + \sin\theta . b,$$

$$L\,b = -\sin\theta . a + \cos\theta . b.$$

We have found incidentally that the possible real *eigenvalues* $\pm 1$ and the non real eigenvalues $e^{i\theta}$ just considered *all have absolute value* 1. This is easy to show directly: $Lz = \lambda z$ implies $L\overline{z} = \overline{\lambda}\,\overline{z}$, hence $(z\,|\,\overline{z}) = (Lz\,|\,L\overline{z}) = |\lambda|^2 (z\,|\,\overline{z})$ and $(z\,|\,\overline{z}) = (a\,|\,a) + (b\,|\,b) \neq 0$.

The above treatment enables us to give the example of totally isotropic subspaces announced on p. 92 . The dimension of $E_Q$ being equal to 2k, the eigenvectors to the eigenvalue $e^{i\theta}$ (or $e^{-i\theta}$) in the complexified space $E + iE$ form a totally isotropic subspace of dimension k.

Summing up: E is the direct sum of mutually orthogonal invariant lines (see p. 102 ) and invariant and irreducible two-planes. These two-planes undergo rotations, the angle being the same in all the two-planes which belong to the same quadratic minimal polynomial.

3.6 The careful discussion of the isometries of Euclidean space will now be amply rewarded. It will serve as a guide and background for the more involved treatment of the corresponding problem in the Lorentz case.

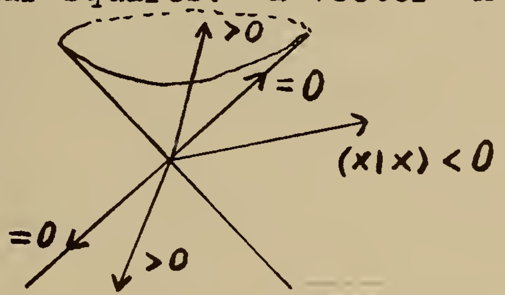
We consider an n-dimensional vector space E over the real field provided with Lorentz metric

(3.6.1) $(x \mid y) = x_1 y_1 - x_2 y_2 - \dots - x_n y_n$ ,

of formula (3.3.4). The corresponding quadratic form

(3.6.2) $(x \mid x) = x_1^2 - x_2^2 - \dots - x_n^2$,

which is the scalar square of the vector x, is indefinite. We classify the non zero vectors $x \in E$ into three classes according to the signs of their scalar squares: a vector x will be called



time-like if $(x \mid x) > 0$,

light-like or isotropic if $(x \mid x) = 0$

space-like if $(x \mid x) < 0$.

The equation $(x \mid x)\ 0$ defines the light cone. Time-like vectors lie inside, space-like vectors outside, the light-cone, whereas light-like vectors lie along its generators.

A time-like or light-like vector x is called positive if $x_1 > 0$ and negative if $x_1 < 0$.

Now we are going to give an axiomatic definition of a Lorentz space. The figure shows there exist vectors inside the cone and that a two-plane through the origin either lies completely outside the light-cone (except at 0), or is tangent to the light cone or cuts the cone. In each case it contains vectors lying outside the cone, that is space-like vectors. This leads us to the following two axioms for the Lorentz metric:

(I) *There exists an* $a \in E$, *such that* $(a \mid a) > 0$, *i.e. there exists at least one time-like vector*.

(II) *Every two-plane contains at least one space-like vector.*

From these axioms the following properties can be deduced.

(A) *If* $a$ *is time-like*, $b \neq 0$ *and* $(a \mid b) = 0$, *then* $b$ *is space-like, i.e.* $(b \mid b) < 0$.

*Proof*. Since $(a \mid b) = 0$ and $(a \mid a) \neq 0$ it follows that $b$ is not collinear with $a$ and thus $a$ and $b$ span a two-plane which by (II) contains a space-like vector $\lambda a + \mu b$. Since $0 > (\lambda a + \mu b \mid \lambda a + \mu b) = \lambda^2 (a \mid a) + \mu^2 (b \mid b)$ and $(a \mid a) > 0$, we have $(b \mid b) < 0$.

(B) *If* $a$ *is light-like and* $(a \mid b) = 0$, *then either* $b$ *is space-like or* $b = \nu a$ (*i.e.* $a$ *and* $b$ *are collinear).*

*Proof*. If $b \neq \nu a$, the manifold spanned by $a$ and $b$ is two-dimensional and thus by axiom (II) there exists a vector $\lambda a + \mu b$ such that $0 > (\lambda a + \mu b \mid \lambda a + \mu b) = \mu^2 (b \mid b)$. Hence $(b \mid b) < 0$.

We state as a corollary:

(C) *Two vectors, either of them time-like or light-like, are never orthogonal to each other, except when both are light-like and collinear*.

It follows that *a Lorentz space is regular*. In fact if $x$ is space-like, then $(x \mid x) \neq 0$; if $x$ is time-like or light-like, then for any time-like $y$ we have $(x \mid y) \neq 0$.

As an other application (outside the scope of the present lectures) we mention, that if $a$ is time-like, then $(a + ib \mid a + ib) \neq 0$ for any $b$. In fact $(a + ib \mid a + ib) = (a \mid a) - (b \mid b) + 2i(a \mid b)$.

If $(a \mid b) \neq 0$, then the imaginary part is $\neq 0$; if $(a \mid b) = 0$, then $(b \mid b) < 0$ and $(a \mid a) - (b \mid b) > 0$. This fact underlies the definition of a hyperbolic differential operator $P\left(\frac{\partial}{\partial x}\right)$ given by Gårding. Restricting ourselves to a homogeneous polynomial $P(x)$, we say that $P$ is hyperbolic if there exists a vector $a$ such that $P(a + ib) \neq 0$ for all real vectors $b$.

We must still show that the axioms (I) and (II) fix the Lorentz signature of the space, which means that with a suitable basis the metric can be written in the form (3.6.1). Let $e_1$ be a a unitary time-like vector, i.e. $(e_1 \mid e_1) = 1$. By (A) the orthogonal complement $H$ of the one-dimensional subspace spanned by $e_1$ is space-like. The restriction of the metric to $H$ is definite negative, and thus, with an obvious modification of sign, it follows from 3.5 that there exists in $H$ a basis $e_2, \ldots, e_n$ such that for $x' \in H$ we have $(x' \mid x') = -\sum_{k=2}^{n} x_k^2$. If $x$ is any vector in $E$, we can write $x = x_1 e_1 + x'$ with $x' = \sum_{k=2}^{n} x_k e_k \in H$ and so

$$(x \mid x) = x_1^2 (e_1 \mid e_1) + (x' \mid x') = x_1^2 - x_2^2 - \ldots - x_n^2.$$

A Lorentz space $E$ has three kinds of non zero subspaces $H$:

1) $H$ is called time-like if it contains some time-like vector;

2) $H$ is called space-like if its non zero vectors are all space-like.

3) $H$ is called light-like if it contains some light-like vector but no time-like vector.

It is clear that all possibilities are hereby exhausted.

We will prove the following: A light-like subspace H contains exactly one light-like line g. The only vectors of H not in g are space-like vectors orthogonal to g.

Proof. By definition H contains a light-like line g. If H contains some vector b not in g then the two-plane spanned by b and a vector $\ell \in g$ contains no time-like vector, i.e. we have

$$0 \geqq (\lambda\ell + b \mid \lambda\ell + b) = 2\lambda(\ell \mid b) + (b \mid b)$$

for all $\lambda$. Thus $(\ell \mid b) = 0$, which, combined with $b \notin g$, implies by (B) that b is space-like.

In the opposite direction we prove that if a subspace H contains exactly one light-like line g then H is light-like. We must prove that H contains no time-like vector. Suppose H contained a time-like vector a, i.e. $(a \mid a) > 0$. Then by (B) we have $(a \mid \ell) \neq 0$ for $\ell \in g$ and therefore $(a + \lambda\ell \mid a + \lambda\ell) = (a \mid a) + 2\lambda(a \mid \ell) = 0$ has a solution in $\lambda$. Thus $a + \lambda\ell$ would be a light-like vector not in g.

Here are some further comments on the definitions of the various subspaces of a Lorentz space. First of all we note the heterogeneous character of the three definitions.

A Lorentz space itself is always time-like; every time-like subspace of a Lorentz space is itself a Lorentz space. In particular a time-like subspace of dimension $\geqq 2$ contains vectors of all three types.

The intersection of a space-like subspace and the light-cone reduces to the vertex of the cone; this property is characteristic for the space-like subspaces. A space-like subspace is obviously isomorphic to a Euclidean space with $(x' \mid x') = -(x \mid x)$.

Finally we describe geometrically the light-like subspaces. Every generator $g$ of the light-cone is a one-dimensional, minimal light-like subspace. The $(n - 1)$ - dimensional tangent plane $T_g$ to the light-cone along the generator $g$ is a maximal light-like subspace, $g$ is the radical of $T_g$, furthermore $T_g = g^{\perp}$, $g = T_g^{\perp}$. Any other light-like subspace $H$ contains some generator $g$ and is contained in the corresponding $T_g$:

$$g \subset H \subset T_g .$$

Conversely this characterizes the light-like subspaces $H$. All this follows easily from the above two theorems on light-like subspaces.

Time-like and space-like subspaces are obviously *regular,* whereas light-like subspaces are *singular,* since $g$ is orthogonal to everything. We emphasize the important fact that *the only totally isotropic subspaces are the light-like lines* $g$.

*The orthogonal complement of a time-like subspace is space-like, that of a space-like subspace is time-like.* The first assertion follows from (A). To prove the second, let $H$ be space-like. Then $H^{\perp}$ is regular (cf. p. 95) and thus either space-like or time-like. But in the first case $E$, which is the direct sum of $H$ and $H^{\perp}$, would also be space-like.

*The orthogonal complement of light-like subspace* $H$ *such that* $g \subset H \subset T_g$ *is a light-like subspace* $H^{\perp}$ *such that* $g \subset H^{\perp} \subset T_g$. *We have the relations* $H \cap H^{\perp} = g$, $H + H^{\perp} = T_g$; the second follows from a general property of orthogonal complements proved on p. 95.

3.7 Let $E$ be a real space with an isometry $L$. The minimal polynomial $m(t)$ of $E$ relative to $L$ is the product of powers of distinct real linear and irreducible quadratic polynomials of the type

$$(t - \varepsilon)^a, \quad \varepsilon^2 = 1, \qquad (t - \lambda)^b, \quad \lambda \neq 0, \qquad \lambda \neq \pm 1,$$

$$[Q(t)]^c, \quad Q(t) = t^2 + \beta t + \gamma, \quad \beta^2 < 4\gamma.$$

The restriction $\lambda \neq 0$ follows from the fact that no factor of the type $t^c$ can occur, since $L$ is one-to-one. We shall always tacitly assume $\lambda \neq 0$ in the future. According to the theory developed in 3.2, $E$ is the direct sum of subspaces denoted by $E_\varepsilon$, $E_\lambda$, $E_Q$ whose minimal polynomials are the above factors of $m(t)$. The particular structure of Lorentz space permits the following precise statement about the factorization of $m(t)$ and the corresponding subspaces.

The minimal polynomial $m(t)$ relative to a isometry $L$ of a Lorentz space $E$ of dimension $n$ can be decomposed into

$$(3.7.1) \qquad m(t) = (t - 1)^p \, (t + 1)^q \, (t - \lambda)^r \, (t - \lambda^{-1})^r \prod_{k=1}^{h} Q_k(t)$$

where $\lambda$ is a real number such that $\lambda^2 \neq 1$, the $Q_k(t)$ are distinct real self-reciprocal irreducible quadratic polynomials with leading coefficient 1. The possible values of $p, q, r, h$ are:

$$(3.7.2) \qquad p = 0,1,3; \quad q = 0,1,3; \quad r = 0,1; \quad h = 0,1,2,\ldots$$

with the following restrictions:

a) $p + q + 2r + 2h \leqq n$,

b) $p, q, r$ can not vanish sumultaneously,

c) the combinations $\{r = 1, p = 3\}$, $\{r = 1, q = 3\}$, $\{p = 3, q = 3\}$ can not occur.

Among the subspaces $E_\varepsilon$, $E_Q$, $E_\lambda, E_{\lambda^{-1}}$, into which $E$ is decomposed, any one of the types $E_\varepsilon$ or $E_Q$ is orthogonal to every subspace except itself, while a subspace of the type $E_\lambda$ is orthogonal to every subspace except $E_{\lambda^{-1}}$. The $E_\varepsilon$ and $E_Q$ are regular; $E_\lambda$ and $E_{\lambda^{-1}}$ are singular, but the direct sum $P$ of $E_\lambda$ and $E_{\lambda^{-1}}$ is regular.

Every $E_Q$ is space-like while the direct sum $P$ of $E_\lambda$ and $E_{\lambda^{-1}}$ is time-like. An $E_\varepsilon$ with minimal polynomial $t - \varepsilon$ may be either time-like or space-like. An $E_\varepsilon$ with minimal polynomial $(t - \varepsilon)^3$ is always time-like. All three of the two subspaces $E_\varepsilon$ and the unique subspace of the kind $P$ cannot be absent simultaneously; moreover exactly one among those present is time-like. These assertions include and motivate the restrictions b) and c).

We turn to the proof of this somewhat long statement. By Example I. for Lemma IV. we know that a subspace $E_\lambda$ whose minimal polynomial is $(t - \lambda)^r$, is totally isotropic. On the other hand, since the dimension of this subspace is $\geqslant r$ and since the only totally isotropic subspaces in a Lorentz space are the light-like lines, we obtain $r = 1$.

It follows in a similar way from Example II. for Lemma IV, that $Q(t) = t^2 + \beta t + \gamma$ must be self-reciprocal, that is $\gamma = 1$. This being the case, Example II. for Lemma VI proves that the exponent of $Q(t)$ in $m(t)$ must be 1.

It follows from Example I. for Lemma VI. that the exponent of $(t - \varepsilon)$ in $m(t)$ must be $\leqq 3$.

Now Lemma III$^{\text{bis}}$ proves that every $E_Q$ is orthogonal to $E_\varepsilon$, $E_\lambda$ and to all the other subspaces of type $E_Q$; $E_\varepsilon$ is orthogonal to $E_\lambda$, to $E_{-\varepsilon}$ and to $E_Q$; $E_\lambda$ is orthogonal to every subspace except possibly to $E_{\lambda^{-1}}$. From the last statement we infer that $(t - \lambda)$ and $(t - \lambda^{-1})$ occur simultaneously, since otherwise the whole space $E$ would be singular.

Furthermore, $E_\lambda$ and $E_\mu$ with $\mu \neq \lambda^{\pm 1}$ cannot occur simultaneously, since both would contain isotropic vectors, which are orthogonal to each other without being collinear, in contradiction to (B).

Finally all subspaces $E_Q$, $E_\varepsilon$, $P = E_\lambda + E_{\lambda^{-1}}$ are regular, since each of them is orthogonal to all others, and on the other hand the entire space $E$ is regular. The light-like subspaces $E_\lambda$ and $E_{\lambda^{-1}}$ are clearly singular.

As to the assertions a), b), c) in the statement, the first is obvious, while the other two will be proved as we go along, together with the assertion at the end of the statement.

3.8 Next we prove that the subspaces $E_Q$ are space-like. Let $a$ be an arbitrary non zero vector of $E_Q$ and define the vector $b$ by (3.5.3). Then $(a \mid b) = 0$ and $(a \mid a) = (b \mid b)$ by (3.5.4) and

(3.5.5). Moreover, because of (3.5.2)

(3.8.1) $2 \sin\theta . b = (L - L^{-1})a = 2(L - (L + L^{-1}) )a = 2(La - \cos\theta . a)$.

The vectors a and La are not collinear, because otherwise the minimal polynomial of a would be linear. This is impossible, since it must divide Q(t). Hence, on account of (3.8.1), a and b are not collinear either and thus span a two-plane. If $u = \alpha a + \beta b$ is an arbitrary non zero vector in this two-plane, then $(u \mid u) = (\alpha^2 + \beta^2)\ (a \mid a)$. Hence sgn $(u \mid u)$ is independent of u. Then, by axiom (II), sgn $(u \mid u) = -1$, and, in particular, sgn $(a \mid a) = -1$, i.e. a is space-like. Since a is an arbitrary vector of $E_Q$, this subspace is also space-like.

The direct sum of the subspaces $E_Q$, belonging to different polynomials Q, is space-like, since the $E_Q$ are mutually orthogonal. It follows that in (3.7.1) the exponents p, q and $\tau$ can not vanish simultaneously, since otherwise E would be identical to the above direct sum and so E itself would be space-like.

The restriction of the metric form to a subspace $E_Q$ is definite negative and it is easy to check that in spite of this difference of sign all the considerations of 3.5 apply; $E_Q$ is the direct sum of two-planes in which L induces a rotation, which we will now call an elliptic rotation, and which is given by the real formulas (3.5.8) or the complex formulas (3.5.9).

3.9 Now suppose that the factors $(t - \lambda)$ and $(t - \lambda^{-1})$, which we know must occur simultaneously, actually occur in the factorization (3.7.1) of m(t), i.e. $\tau = 1$. We already know from 3.7 that the corresponding subspaces $E_\lambda$ and $E_{\lambda^{-1}}$ are light-like lines annihilated by $L - \lambda$ and $L - \lambda^{-1}$ respectively. These lines are different, since $\lambda^2 \neq 1$, i.e. $\lambda \neq \lambda^{-1}$. We shall show that the two-plane P, the direct sum of $E_\lambda$ and $E_{\lambda^{-1}}$, is time-like. Indeed, let $\ell$ and m be any two non zero vectors, $\ell \in E_\lambda$, $m \in E_{\lambda^{-1}}$.

Since $(\ell|\ell) = (m|m) = 0$, it follows that

(3.9.1) $\quad (\ell \pm m | \ell \pm m) = \pm 2 (\ell | m)$.

Since $(\ell|m) \neq 0$ according to (C), one of the two vectors $\ell \pm m$ is time-like, the other space-like, and hence P is time-like.

We have

(3.9.2) $\quad L\ell = \lambda \ell$, $Lm = \lambda^{-1} m$, $\quad \lambda^2 \neq 1$.

Any non zero vector a in P, which is neither in $E_\lambda$ nor in $E_{\lambda^{-1}}$ can be uniquely expressed in the form $a = \ell + m$, $\ell \in E_\lambda$, $m \in E_{\lambda^{-1}}$, $\ell \neq 0$, $m \neq 0$. We set $b = \ell - m$ and note that, by means of the relations

(3.9.3) $\quad a = \ell + m$, $b = \ell - m$, $\quad \ell \in E_\lambda$, $m \in E_{\lambda^{-1}}$,

a and b determine each other in a unique way and that moreover the correspondence is involutory. Indeed (3.9.3) can also be written $b = \ell + (-m)$, $a = \ell - (-m)$ with $\ell \in E_\lambda$, $(-m) \in E_{\lambda^{-1}}$.

We have, according to (3.9.3) and (3.9.1),

(3.9.4) $\quad (a|b) = 0$, $(b|b) = -(a|a) \neq 0$,

the last relation being a direct transcription of (3.9.1).

Of the two cases $\lambda > 0$ and $\lambda < 0$, we first consider the more important one, viz. $\lambda > 0$, and set $\lambda = e^\theta$. Then (3.9.2) becomes

(3.9.5) $\quad L\ell = e^\theta \ell$, $\quad Lm = e^{-\theta} m$, $\qquad \theta \neq 0$.

From this and (3.9.3) we have

$$La = L(\ell + m) = e^\theta \ell + e^{-\theta} m = e^\theta \frac{a+b}{2} + e^{-\theta} \frac{a-b}{2}.$$

This gives for La and, by symmetry, for Lb

(3.9.6)
$$\begin{cases} La = \operatorname{ch}\theta . a + \operatorname{sh}\theta . b & L^{-1} a = \operatorname{ch}\theta . a - \operatorname{sh}\theta . b, \\ Lb = \operatorname{sh}\theta . a + \operatorname{ch}\theta . b & L^{-1} b = -\operatorname{sh}\theta . a + \operatorname{ch}\theta . b, \\ \qquad (a|a) = -(b|b), \quad a \perp b. & \end{cases}$$

The formulae (3.9.5) and (3.9.6) which are connected by (3.9.3) are clearly equivalent. They give the restriction of L to the two-plane P. The corresponding transformation is called a *hyperbolic rotation* in the plane P. Historically speaking, it is the *genuine Lorentz transformation*.

We see that the plane P is *invariant* and *irreducible* under L.

If $\lambda < 0$ we set $\lambda = -e^{\theta}$. The corresponding L appears as the product (taken in any order) of a hyperbolic rotation and a symmetry with respect to the origin. In relativistic terminology such a tranformation is called improper since it interchanges the two halves of the light-cone. The two-plane P is again *invariant* and *irreducible*.

3.10 Finally we focus our attention on one of the spaces $E_{\varepsilon}$, where $\varepsilon^2 = 1$. We suppose, of course, that a corresponding factor actually occurs in (3.7.1). We already know that E is *regular*.

The following long *statement* gives detailed information concerning the structure of the subspaces $E_{\varepsilon}$.

Let $E'_{\varepsilon}$ be that subspace of $E_{\varepsilon}$ which is annihilated by $L - \varepsilon$. $E'_{\varepsilon}$ is invariant and may be regular or singular.

a) If $E'_{\varepsilon}$ is *regular*, $E_{\varepsilon} = E'_{\varepsilon}$ and the minimal polynomial of $E_{\varepsilon}$ is $t - \varepsilon$.

b) If $E'_{\varepsilon}$ is *singular*, $E'_{\varepsilon}$ is a proper *light-like* subspace of $E_{\varepsilon}$, the later being *time-like*. $E'_{\varepsilon}$ contains a unique light-like line g, to which it is orthogonal. The minimal polynomial of $E_{\varepsilon}$ is $(t - \varepsilon)^3$. The subspace $E''_{\varepsilon}$ annihilated by $(L - \varepsilon)^2$ is singular. It is identical to the hyperplane $T_g$ tangent to the light-cone along the generator g. In the succesion $E_{\varepsilon} \supset E''_{\varepsilon} \supset E'_{\varepsilon}$ the dimension decreases by 1 at each step. We have

$$(3.10.1) \qquad (L - \varepsilon)\, E''_{\varepsilon} = g, \quad (L - \varepsilon)^{2} E_{\varepsilon} = g.$$

The subspace $\Pi$ defined by $(L - \varepsilon)\, E_{\varepsilon} = \Pi$ is a light-like two-plane $\Pi \supset g$, $\Pi \perp g$. Furthermore

$$(3.10.2) \qquad E'^{\perp}_{\varepsilon} = \Pi, \qquad \Pi^{\perp} = E'_{\varepsilon},$$

the complements being taken with respect to $E_{\varepsilon}$, and

$$(3.10.3) \qquad E'_{\varepsilon} + \Pi = T_g, \quad E'_{\varepsilon} \cap \Pi = g.$$

The subspaces $T_g = E''_{\varepsilon}$, $E'_{\varepsilon}$, $\Pi$, $g$ are invariant.

Proof and Comments. (a) The fact that $E_{\varepsilon} = E'_{\varepsilon}$, if the latter space is regular, is contained in Lemma VII (p. 99 ). In this case $(L - \varepsilon)$ annihilates the whole space $E_{\varepsilon}$, i.e. $E_{\varepsilon}$ is point-wise invariant if $\varepsilon = 1$, and point-wise reflected in the origin if $\varepsilon = -1$. If $\dim E_{\varepsilon} \geq 2$, the space $E_{\varepsilon}$ can be decomposed into the direct sum of → mutually orthogonal invariant lines in infinitely many ways. These lines are all space-like, if $E_{\varepsilon}$ is time-like (cf. pp. 102 and 107 ).

(b) From now on we suppose that $E'_{\varepsilon}$ is singular, or, in other terms, light-like. Then $E_{\varepsilon}$ is time-like, since it is regular and contains the light-like subspace $E'_{\varepsilon}$ (p. 107 - 108 ). We know (cf. p. 112 ) that the remaining possibilities for the minimal polynomial of $E_{\varepsilon}$ are $(t - \varepsilon)^2$ and $(t - \varepsilon)^3$. We shall show that $(t - \varepsilon)^2$ is impossible.

We first observe that (3.10.2) follows from Lemma VIII$^{\text{bis}}$, with $f(t) = t - \varepsilon$, and that (3.10.3) follows from the theorem at the end of section 3.6, since $g$ is the radical of $T_g$.

Next we prove the first of the equations (3.10.1) stating that $(L - \varepsilon)\, E''_{\varepsilon} = g$. Set $(L - \varepsilon)\, E''_{\varepsilon} = G$. Then $G \subset (L - \varepsilon) E_{\varepsilon} = \Pi$; $(L - \varepsilon) G = (L - \varepsilon)^{2} E''_{\varepsilon} = 0$, that is $G \subset E'_{\varepsilon}$. Hence $G \subset \Pi \cap E'_{\varepsilon} = g$. On the other hand let $u \in g$. Then $u \in \Pi = (L - \varepsilon) E_{\varepsilon}$. Therefore there exists $y \in E_{\varepsilon}$ such that $(L - \varepsilon) y = u$. Now $(L - \varepsilon)^{2} y = (L - \varepsilon)\, u = 0$, By a curious misfortune a long and distorting omission occurred in the sentence of lines 13 - 15 on p. 116. You may either delete this vicious sentence or complete it. Both are easy.

M. Riesz

since $u \in g \subset E'_{\varepsilon}$. Consequently $y \in E''_{\varepsilon}$ and $u = (L - \varepsilon)y \in (L - \varepsilon) E''_{\varepsilon} = G$, that is $g \subset G$. Summing up, we have $G = g$.

Now we prove that $E''_{\varepsilon} = T_g$ and begin by showing that $E''_{\varepsilon} \subset T_g$. Let $y \in E''_{\varepsilon}$ and set $(L - \varepsilon)y = u$. Then, as just shown, $u \in g$. Now, if $u = 0$, $y \in E'_{\varepsilon} \subset T_g$. If $u \neq 0$ we first note that $u \in g$ implies $(u \mid u) = 0$ and secondly we use the isometry relation $0 = (Ly \mid Ly) - (y \mid y)$, which in view of $Ly = \varepsilon y + u$, $\varepsilon^2 = 1$ and $(u \mid u) = 0$ reduces to $0 = 2\varepsilon(y \mid u)$ or $(y \mid u) = 0$. From $u \in g$ and $u \neq 0$ it then follows that $y \perp g$, that is $y \in T_g$.

On the other hand, the assertion $T_g \subset E''_{\varepsilon}$ is clearly equivalent to the assertion $(L - \varepsilon)^2 T_g = 0$, which we shall show to be valid. We know that $T_g = E'_{\varepsilon} + \Pi$. Now $(L - \varepsilon)^2 E'_{\varepsilon} = 0$, since $(L - \varepsilon)E'_{\varepsilon} = 0$, and $(L - \varepsilon)^2 \Pi = (L - \varepsilon)^3 E_{\varepsilon} = 0$, since as proved on p. 112 , the exponent p of the minimal polynomial $(t - \varepsilon)^p$ of E is $\leq 3$.

Herewith the proof of the assertion $E'' = T_g$ is completed.

Now we are able to prove that the minimal polynomial of $E_{\varepsilon}$ is $(t - \varepsilon)^3$. Since $E_{\varepsilon}$ is regular and $E''_{\varepsilon} = T_g \subset E_{\varepsilon}$ is singular, $E''_{\varepsilon}$ is a proper subspace of $E_{\varepsilon}$. Thus $(L - \varepsilon)^2$ does not annihilate the entire space $E_{\varepsilon}$ and therefore the minimal polynomial of $E_{\varepsilon}$ is $(t - \varepsilon)^3$.

Before going further, we collect some facts which are obvious by now. The minimal polynomial of $E_{\varepsilon}$ is $(t - \varepsilon)^3$, that of $E''_{\varepsilon} = T_g$ and of $\Pi$ is $(t - \varepsilon)^2$, that of $E'_{\varepsilon}$ and of g is $t - \varepsilon$. The minimal polynomial of x in $E_{\varepsilon}$ but not in $T_g$ is $(t - \varepsilon)^3$; important examples of such x are any time-like vector or any light-like vector which is not in g. The minimal polynomial of y in $T_g$ but not in $E'_{\varepsilon}$ , or equivalently, in $\Pi$ but not in g, is $(t - \varepsilon)^2$; that of v in E' but not $= 0$ is $(t - \varepsilon)$.

The second of the equations (3.10.1) stating that $(L - \varepsilon)^2 E_\varepsilon = g$ follows by Lemma $\text{VIII}^{\text{bis}}$ from $g = T_g^{\perp}$ and $T_g = E''_\varepsilon$.

Now we arrive at the proof of one of the main assertions in our statement, namely that $\Pi$ is a light-like two-plane. The light-likeness follows from $g \subset \Pi \subset T_g$. We recall that $g$ is a proper subspace of $\Pi$, since $(L - \varepsilon)\, g = 0$ and $(L - \varepsilon)\Pi = g \neq 0$. Let $y \in \Pi$, $\tilde{y} \in \Pi$, $\tilde{y} \notin g$, hence $\tilde{y} \notin E'_\varepsilon$. Then $(L - \varepsilon)y = z \in g$, $(L - \varepsilon)\tilde{y} = \tilde{z} \in g$, $\tilde{z} \neq 0$. Hence $z = \lambda \tilde{z}$. Set $(y - \lambda \tilde{y}) = v$. Then $v \in \Pi$. Moreover $(L - \varepsilon)\, v = z - \lambda \tilde{z} = 0$. Thus $v \in E'_\varepsilon$. Consequently $v \in \Pi \cap E'_\varepsilon = g$, that is $v = \mu \tilde{z}$. Finally $y = \lambda \tilde{y} + \mu \tilde{z}$, hence $\dim \Pi = 2$.

Clearly we have $\dim T_g = \dim E_\varepsilon - 1$. Furthermore the relation $\dim E'_\varepsilon + \dim \Pi = \dim E_\varepsilon$ ( cf. pp. 94 - 95 ) implies $\dim E'_\varepsilon = \dim E_\varepsilon - 2$. These facts also follow from the equation $(\tilde{z} \mid x) = 0$ for $T_g$ and the equations $(\tilde{z} \mid x) = 0$ and $(\tilde{y} \mid x) = 0$ for $E'_\varepsilon$.

3.11 We know that the only possible exponent $> 1$ with which $(t - \varepsilon)$ can occur in the minimal polynomial of a Lorentz transformation $L$ is 3, and we just have given a detailed description of an invariant subspace $E_\varepsilon$, whose minimal polynomial is $(t - \varepsilon)^3$. However, we still do not know whether such a subspace actually can occur. We shall now construct a Lorentz transformation whose minimal polynomial is $(t - \varepsilon)^3$.

Such a transformation is called a parabolic rotation or reflection, for reasons to appear later.

The light-like two-plane $\Pi$ of 3.10 will play a fundamental role in our construction as revealed by the following theorem.

Theorem. Let $E$ be a Lorentz space of dimension $m \geqq 3$ and $\Pi$ a light-like two-plane in $E$. There exist isometries $L$ of $E$ such that $(L - \varepsilon)E = \Pi$. The minimal polynomial of $E$ relative to such an $L$ is $(t - \varepsilon)^3$. In the case of $\varepsilon = 1$, the isometries $L$, in the case of $\varepsilon = -1$, the isometries $(-L)$ - in both cases completed by the identity- form a one-parameter Abelian group.

We first analyze the implications of the assumption that $L$ is an isometry and that $(L - \varepsilon) E = \Pi$. The light-like line in $\Pi$ will, as usual, be denoted by $g$, tangent $(m - 1)$- plane to the light cone in $E$ along $g$ by $T_g$, and the orthogonal complement of $\Pi$ with respect to $E$ by $E'$. As we know (cf. p. 109 ) $\Pi \cap E' = g$, $\Pi + E' = T_g$. All these proper subspaces of $E$ are singular, and all are determined by $\Pi$, since $g = \operatorname{rad} \Pi$, $T_g = g^{\perp}$, $E' = \Pi^{\perp}$.

From $(L - \varepsilon) E = \Pi$ and $\Pi^{\perp} = E'$, it follows by Lemma $\text{VIII}^{\text{bis}}$ that $E'$ is the subspace of $E$ and $g = \Pi \cap E'$ the subspace of $\Pi$ which are annihilated by $L - \varepsilon$. Next we prove that $(L - \varepsilon) \Pi = g$. To this end let $0 \neq z \in g$. We first show that there exists a $y \in \Pi$ such that $(L - \varepsilon)y = z$. Since $(L - \varepsilon)E = \Pi$ and $z \in g \subset \Pi$, there exists a $v \in E$ such that $(L - \varepsilon)v = z$, or equivalently $Lv = \varepsilon v + z$. The isometry relation $0 = (Lv \mid Lv) - (v \mid v)$ together with $\varepsilon^2 = 1$ and $(z \mid z) = 0$ implies that $0 = 2\varepsilon(v \mid z) + (z \mid z) = 2\varepsilon(v \mid z)$ and then $(v \mid z) = 0$. Hence $v \perp g$ or $v \in g^{\perp} = T_g = \Pi + E'$, that is $v = y + u$, $y \in \Pi$, $u \in E'$. Then $z = (L - \varepsilon)v = (L - \varepsilon)y + (L - \varepsilon)u = (L - \varepsilon)y$, as we had desired to show. Since $(L - \varepsilon)y = z \neq 0$, $(L - \varepsilon)z \in (L - \varepsilon)g = 0$, $y$ and $z$ are not collinear, and hence they form a basis for $\Pi$, $\Pi = \{\lambda y + \mu z\}$. Consequently, $(L - \varepsilon)\Pi = \{\lambda z\} = g$, as asserted above.

Summing up, we have as consequences of our assumptions

$$(L - \varepsilon)E = \Pi;\ (L - \varepsilon)^2 E = (L - \varepsilon)\Pi = g;\ (L - \varepsilon)^3 E = (L - \varepsilon)^2 \Pi = (L - \varepsilon)g = 0. \tag{3.11.1}$$

This tells us that the respective minimal polynomials of $E$, $\Pi$, $g$ are $(t - \varepsilon)^3$, $(t - \varepsilon)^2$, $(t - \varepsilon)$.

We have shown that to every $0 \neq z \in g$ there exists a $y \in \Pi$ such that $(L - \varepsilon)y = z$. Clearly, this $y$ is not unique, since the above relation implies that $(L - \varepsilon)(y + \mu z) = z$ for an arbitrary $\mu$. None of these $y$ belongs to $g$, since $(L - \varepsilon)g = 0$. On the other hand, as we have shown a while ago, $u \in \Pi$ but $u \notin g$

implies that $(L - \varepsilon)\, u \neq 0$. Taking all this together, we see that the different $y \in \Pi$ satisfying $(L - \varepsilon)\, y = z$, with a given non zero $z \in g$ form a line parallel to g but different from g.

Having gone through these details, select a $y \in \Pi$, $y \notin g$ and set $(L - \varepsilon)\, y = z$. Then $z \neq 0$, $z \in g$, $(z \mid z) = 0$, while $y \in \Pi$ implies that $y \perp g$. Since $y \notin g$, y is space-like, that is $(y \mid y) < 0$. Collecting these facts, we may state: From $(L - \varepsilon)\, E = \Pi$, $y \in \Pi$, $y \notin g$, $(L - \varepsilon)\, y = z$ it follows that

$$(3.11.2) \quad (z \mid z) = 0,\ z \in g,\ z \neq 0,\ (y \mid z) = 0,\ y \perp g,\ (y \mid y) < 0.$$

We claim that *for an arbitrary fixed pair* $y, z$, $y \in \Pi$, $y \notin g$, $z \in g$, $z \neq 0$, *there exists one and only one isometry L of E such that* $(L - \varepsilon)\, E = \Pi$ *and* $(L - \varepsilon)\, y = z$. In order to fulfill the requirement $(L - \varepsilon)\, E = \Pi$, we seek a vector $x \in E$ and an isometry L such that $(L - \varepsilon)\, x = y$. With the further requirement $(L - \varepsilon)\, y = z \neq 0$ and its consequence $(L - \varepsilon)\, z = 0$ we have

$$(3.11.3) \quad (L - \varepsilon)\, x = y,\ (L - \varepsilon)\, y = z \neq 0,\ (L - \varepsilon)\, z = 0,$$

or, equivalently,

$$(3.11.4) \quad Lx = \varepsilon x + y,\quad Ly = \varepsilon y + z,\quad Lz = \varepsilon z \neq 0.$$

(In learned terminology we have here a *Jordan normal form corresponding to an elementary divisor of order 3.*

The postulate that L is an isometry requires that

$$(3.11.5) \quad (Ls \mid Lt) = (s \mid t),$$

s and t denoting any two of the vectors x, y, z, distinct or identical. It is readily seen that, as far as the vectors y, z only are concerned, these relations are automatically fulfilled by virtue of (3.11.2), (3.11.4) and $\varepsilon^2 = 1$. Now we will find the requirements which the isometry postulate imposes on the vector x. According to the equations (3.11.4) and $\varepsilon^2 = 1$, the postulates (3.11.5) are respectively equivalent, 1) for $s = x$, $t = z$ to $(y \mid z) = 0$, which is already known by (3.11.2) 2) for $s = x$, $t = y$ to $(x \mid z) + (y \mid y) = 0$, 3) for $s = x$, $t = x$ to $2\varepsilon(x \mid y) + (y \mid y) = 0$. Thus we have as requirements upon x,

(3.11.6) $(x \mid z) = -(y \mid y) > 0, \; \varepsilon(x \mid y) = -\frac{1}{2}(y \mid y) > 0.$

The vectors y, z and any vector x which satisfies (3.11.6) are linearly independent, since $y, z \in \Pi$, $0 \neq z \in g$, $y \notin g$ and, on the other hand, $x \notin \Pi$, because $(x \mid z) \neq 0$, i.e. x is not orthogonal to g. This being the case, we extend the transformation L by linearity to the space spanned by x, y, z and find for any vector v in this space $(Lv \mid Lv) = (v \mid v)$ as a consequence of (3.11.5).

Now let us suppose for a short while that the dimension of the Lorentz space E is equal to 3. Hence E is identical to the space spanned by the vectors x, y, z. The transformation L is then defined by the above in the entire space E, and is an isometry if (3.11.4) and (3.11.6) are fulfilled. Thus the only remaining thing is to find a vector x satisfying equations (3.11.6).

Now the equations (3.11.6) define for x two non homogeneous planes (that is, not passing through the origin). Since y and z are not collinear, these planes intersect in a non homogeneous line h. Any vector x from the origin to a point on h satisfies (3.11.6). Picking any such vector x, the transformation L defined by (3.11.4) and its linear extension is an isometry of the space E. For a generic vector $v = \alpha x + \beta y + \gamma z$ we have by (3.11.3)

$u = (L - \varepsilon) v = \alpha y + \beta z$. Thus $u \in \Pi$, and since y, z are linearly independent, $L - \varepsilon$ maps E onto $\Pi$. Furthermore $u \in g$, if and only if $\alpha = 0$; $u = 0$, if and only if $\alpha = 0$, $\beta = 0$. The construction in the case $m = 3$ is thus completed.

The space E, clearly invariant under L, is also irreducible in the sense of the definition given on p. 96.

We add a few comments concernig the degree of freedom in the choice of the vector x. The line h is parallel to g or, in other terms, the admissible vectors x are congruent mod. g; conversely any such vectors are simultaneously admissible. Indeed, the solution of the corresponding homogeneous equations $(y \mid u) = 0$, $(z \mid u) = 0$ are the vectors $u \in g$. On the other hand, the equations (3.11.4) = (3.11.3) imply that $(L - \mathcal{E})\, g = 0$. Hence any triples $x_1$, y, z and $x_2$, y, z such that $x_1 - x_2 \in g$ satisfy equations (3.11.4) simultaneously. Consequently all linearly extended equations (3.11.4) written with the fixed vectors y, z and any of the $\infty^1$ admissible vectors x yield the same isometry L of the space E.

In the case $\dim E = m > 3$, equations (3.11.6) determine an $(m - 2)$ - dimensional inhomogeneous subspace H. Any vector x from the origin to a point on H satisfies (3.11.6). The isometry L is then defined as above on the three-space $E^{(3)}$ spanned by x, y, z. This subspace is clearly time-like and also invariant and irreducible under L.

It remains to extend the definition of L to the entire space E. Let U be the orthogonal complement of $E^{(3)}$ in E. The subspace U is space-like and E is the direct sum of $E^{(3)}$ and U. Since $E^{(3)} \supset \Pi$, we have $U \subset \Pi^{\perp} = E'$. Now, if there exists an isometry L such that $(L - \mathcal{E})\, E = \Pi$, we must have, according to Lemma $\text{VIII}^{\text{bis}}$, $(L - \mathcal{E})\, E' = 0$. Hence we define the extension of L to U by the relation $(L - \mathcal{E})\, U = 0$. Then U is invariant under L. Furthermore, if we define L on E by linearity, it will be an isometry. Indeed any $w \in E$ can in a unique way be written as $w = v + u$, $v \in E^{(3)}$, $u \in U$, hence $(v \mid u) = 0$. Also $Lv \in E^{(3)}$, $Lu \in U$ and thus $(Lv \mid Lu) = 0$. Taking into account that $(Lv \mid Lv) = (v \mid v)$, $Lu = \varepsilon u$ and $\varepsilon^2 = 1$, we have

$$(Lw \mid Lw) = (Lv + Lu \mid Lv + Lu) = (Lv \mid Lv) + (Lu \mid Lu)$$

$$= (v \mid v) + (u \mid u) = (v + u \mid v + u) = (w \mid w).$$

This shows that L is an isometry for the whole space E. Clearly the minimal polynomial of $\pi$ is $(t - \varepsilon)^2$ and that of E is $(t - \varepsilon)^3$. The subspaces E' and $E'' = T_g$ defined on p. 119 as the orthogonal complements to $\pi$ and g respectively can be written $E' = g + U$, $E'' = T_g = \pi + U$, the sums being orthogonal and direct. We see that E' and E" are those subspaces of E which are annihilated by $(L - \varepsilon)$ and $(L - \varepsilon)^2$ respectively. We also see that $(L - \varepsilon)E = \pi$, $(L - \varepsilon)T_g = g$, $(L - \varepsilon)^2 E = (L - \varepsilon)\pi = g$. All this is independent of the choice of the vector x.

As in the case of $m = 3$, we investigate the degree of freedom in the choice of the vector x, the vectors y and z having been fixed. The nonhomogeneous $(m - 2)$-plane H, defined by the equations (3.11.6) is parallel to E'. This means that all the admissible vectors x are congruent modulo E' and that conversely all such vectors are simultaneously admissible. Indeed, in the present case, the solutions of the corresponding homogeneous equations $(y \mid u) = 0$, $(z \mid u) = 0$ are the vectors $u \in \pi^{\perp} = E'$. On the other hand, because of $(L - \varepsilon)E' = 0$, we obtain, through a reasoning analogous to the one used in the case $m = 3$, the fact that all equations (3.11.4) written with the fixed vectors y, z and any of the $\infty^{m-2}$ admissible vectors x, and extended to the whole space E in the way given above, yield the same isometry L of this space.

The three-space $E^{(3)}$ spanned by the vectors x, y, z, its orthogonal complement U and the decomposition $E = E^{(3)} + U$, all of which were essential tools in our construction, did not appear in the last development, culminating in the uniqueness of an isometry L such that $(L - \varepsilon) E = \Pi$, $(L - \varepsilon) y = z$, $y \in \Pi$, $y \notin g$, $z \in g$, the couple of vectors y, z being fixed. One would expect that all the subspaces $E^{(3)}$ spanned by y, z and any admissible vector x are identical. Indeed, in our construction $E^{(3)}$ appeared as carrying the virulent part of the transformation L, while U was point-wise invariant for $\varepsilon = 1$, and point-wise reflected in the origin, for $\varepsilon = -1$. However, the conjecture put forward above does not agree with the facts, the subspace $E^{(3)}$ and the corresponding decomposition are not unique, as we shall see in a moment.

Let L be the isometry considered above and let $E^{(3)}$ be an arbitrary three-dimensional time-like space such that $E \supset E^{(3)} \supset \Pi$. We shall prove that $E^{(3)}$ and its orthogonal complement U are invariant under L, $(L - \varepsilon) E^{(3)} = \Pi$, $(L - \varepsilon) U = 0$, and that the minimal polynomial of $E^{(3)}$ is $(t - \varepsilon)^3$.

We claim that there exist infinitely many vectors $x \in E^{(3)}$, congruent modulo g, which satisfy equations (3.11.6). To prove this, let $q$ be any line in $E^{(3)}$ which does not belong to $T_g$, i.e. g is not orthogonal to g. Then $(q \cap \Pi) \subset (q \cap T_g) = 0$ and $E^{(3)} = q + \Pi$, the sum being direct. Let $0 \neq v \in q$ ; then $(v \mid z) \neq 0$, according to our assumption. Set $x = \alpha v + \beta y + \gamma z \in E^{(3)}$. The equations (3.11.6) yield

$$(\alpha v + \beta y + \gamma z \mid z) = -(y \mid y), \quad (\alpha v + \beta y + \gamma z \mid y) = -\frac{\varepsilon}{2} (y \mid y).$$

Since $(z \mid z) = (y \mid z) = 0$, this reduces to

$$\alpha (v \mid z) = -(y \mid y), \quad \alpha (v \mid y) + \beta (y \mid y) = -\frac{\varepsilon}{2} (y \mid y).$$

Since $(v \mid z) \neq 0$ and $(y \mid y) \neq 0$ (as we know, $(y \mid y) < 0$), $\alpha$ and $\beta$ can be solved for in a unique way, while $\gamma$ remains arbitrary. The remainder of our assertion should be obvious by now.

We see that to the $\infty^{m-2}$ admissible vectors x these correspond $\infty^{m-3}$ decompositions, those vectors x which are congruent modulo g giving the same decomposition.

By the preceding we see that in the case of dim E $>$ 3, the subspaces $E^{(3)}$ do not play any intrinsic role in the structure of the isometry L. The explanation is the following. For any choice of an admissible x, the stable subspace is not the subspace U defined as the orthogonal complement of the three-space $E^{(3)}$ spanned by the vectors x, y, z, but the subspace $E' = U + g$. This subspace $E' = \Pi^{\perp}$ is determined by $\Pi$ and is entirely independent of the choice of x.

In this connection we consider briefly the particular case of a four-dimensional Lorentz space, i.e. the space-time of special relativity.

Denote the space by E again and let L be a Lorentz transformation in E with the minimal polynomial $(t - \varepsilon)^3$. Setting as before $(L - \varepsilon)E = \Pi$, rad $\Pi = g$, $g^{\perp} = T_g$, we have here dim $T_g = 3$. The subspace $E' = \Pi^{\perp}$ which is point-wise invariant if $\varepsilon = 1$, and point-wise reflected in the origin if $\varepsilon = -1$, is here a light-like two-plane orthogonal to the light-like two-plane $\Pi$ and having with it the generator g in common.

Let $E^{(3)}$ be an arbitrary time-like therree-space such that $\Pi \subset E^{(3)} \subset E$. The orthogonal complement of $E^{(3)}$ is a space-like line $U \subset E'$. Both $E^{(3)}$ and U are invariant under L, the minimal polynomial of $E^{(3)}$ is $(t - \varepsilon)^3$ and that of U is $(t - \varepsilon)$. To obtain the $\infty^1$ decompositions $E = E^{(3)} + U$, it is most convenient to choose the line $U \neq g$ arbitrary in E' and obtain the corresponding $E^{(3)}$ by the relation $E^{(3)} = U^{\perp}$.

For the occurrence of parabolic rotations in the theory of electromagnetic fields see section 4.17.

Now we prove the last part of the theorems stated on p. 119

concerning the group property. Without loss of generality we can suppose $\varepsilon = 1$. We shall show that all isometries such that $(L - 1) E = \Pi$, completed with the identity, form a one-parameter Abelian group. As proved above, any such isometry is entirely determined by the choice of a couple $y \in \Pi$, $y \notin g$, $0 \neq z \in g$ and the equation $(L - 1) y = z$. A second equation $(L - 1) z = 0$ is contained in the equation $(L - 1) g = 0$, which in a general way follows from $(L - 1) E = \Pi$. Now we still fix y, but let z vary along the line g, that is we set $z = \alpha \tilde{z}$, where the vector $0 \neq \tilde{z} \in g$ is fixed and $\alpha$ is a real parameter. If the corresponding isometry is denoted by $L_\alpha$, we have $L_\alpha y = y + \alpha \tilde{z}$, $L_\alpha \tilde{z} = \tilde{z}$. In an analogous way $L_\beta y = y + \beta \tilde{z}$, $L_\beta \tilde{z} = \tilde{z}$. From this it follows that $L_\alpha (L_\beta y) = L_\alpha (y + \beta \tilde{z}) = y + \alpha \tilde{z} + \beta \tilde{z} = y + (\alpha + \beta) \tilde{z}$, $L_\alpha (L_\beta \tilde{z}) = \tilde{z}$. Hence $L_\alpha L_\beta = L_\beta L_\alpha = L_{\alpha + \beta}$. The isometry $L_0$ is clearly the identity.

# Chapter IV

## Representation of isometries by infinitesimal transformations and Clifford bivectors

4.1 In Chapter III we have carried out a thorough investigation of the structure of the isometries of the real Euclidean space and the real Lorentz space. Here we shall reconsider our results in the light of the concept of infinitesimal transformations.

Let $E$ be an n-dimensional vector space provided with a symmetric metric $(x \mid y)$. The transformation $L = 1+\varepsilon T$, where $T$ is a linear transformation, is called an infinitesimal isometry if $(Lx \mid Lx) = (x \mid x)$ is correct in first approximation, that is, if $\varepsilon^2$ is neglected. To this approximation we have

$$(x + \varepsilon Tx \mid x + \varepsilon Tx) = (x \mid x) + 2\varepsilon (x \mid Tx).$$

Hence the necessary and sufficient condition for $L$ to be an infinitesimal isometry is that

$$(x \mid Tx) = 0. \tag{4.1.1}$$

Writing this relation also for $(x \pm y)$, instead of $x$, we obtain

$$(x \mid Ty) + (y \mid Tx) = 0, \tag{4.1.2}$$

which shows that also $(x + \varepsilon Tx \mid y + \varepsilon Ty) = (x \mid y)$ is correct in first approximation. A linear transformation $T$ with the equivalent properties (4.1.1) and (4.1.2) is called skew-symmetric (or shorter: skew) relative to the metric $(x \mid y)$. We show by induction that $T^p$ is skew-symmetric for any odd $p$ and symmetric for any even $p$ or, more briefly,

$$(x \mid T^p y) = (y \mid (-T)^p x). \tag{4.1.3}$$

Indeed if (4.1.3) holds for any $p \leqslant q$, we find by using $T^{q+1} = T^q T = T T^q$, that

$$(x \mid T^{q+1} y) = - (Ty \mid (-T)^q x) = (y \mid (-T)^{q+1} x).$$

If $f(t)$ is a polynomial, we have $(x \mid f(T)y) = (y \mid f(-T)x)$. This result corresponds to Lemma II. in 3.4. The other Lemmas concerning isometries also have their couterparts in the present case of infinitesimal isometries; they are in general simpler in form and easier to prove.

Let not $A$ be any linear transformation of a finite dimensional vector space $E$ into itself. We define a new linear transformation $\exp A = e^A$ by setting

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \tag{4.1.4}$$

We shall prove in section 4.18 that, in an appropriate sense, the right hand side converges like an absolutely convergent numerical series and can be dealt with in a similar manner. Now let $B$ be another linear transformation and suppose that $A$ and $B$ commute with each other. Then we have

$$e^A e^B = e^{A+B}. \tag{4.1.5}$$

Indeed,

$$e^A e^B = \sum_{k=0}^{\infty} \frac{A^k}{k!} \sum_{l=0}^{\infty} \frac{B^l}{l!} = \sum_{m=0}^{\infty} \sum_{k+l=m} \frac{A^k}{k!} \frac{B^l}{l!}$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k+l=m} \binom{m}{k} A^k B^l = \sum_{m=0}^{\infty} \frac{(A+B)^m}{m!} = e^{A+B}.$$

In particular, if $s$ and $t$ are scalars

$$e^{sA}\, e^{tA} = e^{(s+t)A}. \tag{4.1.6}$$

Furthermore we have the following important fact:

If $T$ is skew-symmetric then $e^T$ is an isometry. We obtain by means of (4.1.3) that $(x \mid e^T y) = (y \mid e^{-T} x)$. Replacing $x$ by $e^T x$ and using (4.1.5) it follows that

$$(e^T x \mid e^T y) = (x \mid y). \tag{4.1.7}$$

If $s$ is a continuous real (or complex) parameter, then $sT$ is also skew-symmetric, and $L(s) = e^{sT}$ is a one-parameter Abelian group of isometries, where the identity corresponds to $s = 0$.

We have now according to (4.1.6)

$$\lim_{h \to o} \frac{e^{(s+h)T} - e^{(s-h)T}}{2h} = \lim_{h \to o} e^{sT} \frac{e^{hT} - e^{-hT}}{2h} = e^{sT}.T,$$

because

$$e^{hT} - e^{-hT} = 2hT + 2 \sum_{k=1}^{\infty} \frac{(hT)^{2k+1}}{(2k+1)!} .$$

Hence $de^{sT}/ds = e^{sT}.T$ and by repetition $d^n e^{sT}/ds^n = e^{sT}.T^n$.

In particular

$$\left.\frac{de^{sT}}{ds}\right|_{(s=o)} = T \quad \text{and} \quad \left.\frac{d^n e^{sT}}{ds^n}\right|_{(s=o)} = T^n . \tag{4.1.8}$$

We have just seen that to every skew-symmetric $T$ there belongs a one-parameter group of isometries $e^{sT}$. Conversely, if $L(s)$ is a one-parameter group of isometries, depending analytically on $s$, such that $L(s+t) = L(s)L(t)$, then $L(s) = e^{sT}$ with a skew-symmetric $T$. To prove this, let

$$\lim_{h \to o} \frac{L(s+h) - L(s-h)}{2h} = L(s).\,T,$$

where

$$T = \lim_{h \to o} \frac{L(h)-L(-h)}{2h}.$$

Hence

$$\frac{d}{ds} L\,(s) = L(s).T,$$

and as above

$$\frac{d^n L(s)}{ds^n} = L(s).T^n,$$

and since necessarily $L(0) = 1$,

$$\left.\frac{d^n L(s)}{ds^n}\right|_{(s=o)} = T^n.$$

It follows that

$$L(s) = e^{sT},$$

since both functions are analytic and they have the same derivatives at the origin.

Furthermore $T$ is skew-symmetric, since

$$L(h) - L(-h) = L(h) - L^{-1}(h)$$

and

(4.1.9) $$((L - L^{-1})x \mid x) = (Lx \mid x) - (L^{-1}x \mid x) = 0.$$

In agreement with the terminology used in the theory of Lie groups, we say that $T$ is the infinitesimal transformation of the one-parameter group $L(s)$. This group, as well as any particular element of this group, are said to be generated by the infinitesimal transformation $T$.

Let $G$ be a group continuously depending on a finite number of parameters. Those elements of $G$ which can be connected with the identity by some continuous arc form a subgroup $G_o$, called the connected component of the identity.

Every isometry $e^T$ generated by a skew-symmetric transformation $T$ belongs to the connected component of the identity, since we have a continuous passage by means of the isometries $e^{sT}$ with $s$ varying from 0 to 1.

One of our principal aims is to establish that a perfect converse of this statement is valid as far as isometries of the real Euclidean space and the real Lorentz space are concerned, namely that every isometry of these metrics belonging to the connected component of the identity can be generated by some infinitesimal transformation. This property is the more remarkable, since it is not shared by any other real metric.

4.2. Let $E$ be an n-dimensional real vector space with either a Euclidean or a Lorentz metric. Let $L$ be an isometry of $E$ and let us begin by considering a subspace $E_Q$ of $E$ whose minimal polynomial is the irreducible quadratic polynomial $Q(t) = t^2 + \beta t + 1$, with $|\beta| < 2$; see 3.5 (p. 101) and 3.8 (p. 112). As usual, we also denote the restriction of $L$ to $E_Q$ by the same letter $L$. We set

$$\beta = -2\cos\theta. \tag{4.2.1}$$

Thereby $\sin\theta$ also is determined except for sign. In the following we mean by $\theta$ any solution of equation (4.2.1).

We write the formula (3.5.1) in the form

$$\frac{L+L^{-1}}{2} = \cos\theta \ . \tag{4.2.2}$$

We define the linear transformation $T$ of $E_Q$ into itself by

$$T = \frac{L - L^{-1}}{2\sin\theta} \ . \tag{4.2.3}$$

This formula determines $T$ up to a factor $(-1)$, while $\sin\theta . T$ is entirely determined. According to (4.1.9), $T$ is skew-symmetric; moreover by (3.5.3) and (3.5.7) we get the basic relation

$$T^2 = -1, \tag{4.2.4}$$

which yields

$$e^{\theta T} = \cos\theta + \sin\theta . T. \tag{4.2.5}$$

Combining (4.2.2), (4.2.3), (4.2.5) we find

$$L = \cos\theta + \sin\theta . T = e^{\theta T}, \quad L^{-1} = \cos\theta - \sin\theta . T = e^{-\theta T}. \tag{4.2.6}$$

We are going to analyze the properties and the structure of a skew-symmetric transformation $T$ like the one given by (4.2.3), having the property $T^2 = -1$.

We place ourselves for a while in a more general situation and make the following assumptions: $G$ is a real vector space, regular with respect to a certain metric, $T$ is a skew-symmetric linear transformation of $G$ into itself, $T^2 = -1$ in $G$. Such a linear transformation $T$ will be called a skew-symmetric square root of $(-1)$ in $G$.

We prove that such a transformation $T$ is also an isometry in G. More precisely, any two of the three properties

a) T is skew-symmetric,

b) $T^2 = -1$,

c) T is an isometry,

imply the third.

Proof. If a) and b) are satisfied, then $(Tx \mid Tx) = -(T^2x \mid x) = (-x \mid x) = (x \mid x)$, hence also c) is satisfied. If b) and c) are satisfied, then $(x \mid Ty) = (Tx \mid T^2y) = (Tx \mid -y) = -(Tx \mid y)$, hence a) is satisfied. Finally, if c) and a) satisfied, then $(x \mid y) = (Tx \mid Ty) = -(T^2x \mid y)$, that is $(x + T^2x \mid y) = 0$ for all $x \in G$, $y \in G$. Hence, in view of the regularity of the metric, $x + T^2x = 0$ for every $x \in G$, that is also b) is satisfied.

Next we observe that two subspaces H and $H^{\perp}$ which are orthogonal complements of each other with respect to G are simultaneously invariant under any skew-symmetric transformation T (in G). This follows immediately from the relation $(x \mid Ty) + (y \mid Tx) = 0$.

A skew-symmetric square root of $(-1)$ has no real eigenvectors. Indeed $T^2a = -a$ and $Ta = \lambda a$ gives $\lambda^2 = -1$. Hence dim $G \geqslant 2$. If G is a two-plane, T is one of the two possible rotations by the angle $\frac{\pi}{2}$ in this plane. If dim $G > 2$, G can be decomposed by means of the transformation T into the direct sum of mutually orthogonal two-planes invariant under T, which implies that dim G always is even. To prove our assertion, let $a \neq o$ be an arbitrary vector in G, set $Ta = b$. Since a and b are not collinear, they span a two-plane $B_1$ say. Next take a vector c orthogonal to $B_1$. Then, as just shown, also $Tc = d$ is orthogonal to $B_1$. Hence c and d span a second two-plane $B_2$ orthogonal to $B_1$. The continuation is obvious. In each of the two-planes T produces a rotation by the angle $\frac{\pi}{2}$. Reversing the steps, we have a procedure to construct arbitrary transformations of the type T in G. We decompose G, in an arbitrary way into mutually orthogonal two-planes $B_k$, take one of the two skew-symmetric square roots (-1) in $B_k$, denote it for the sake of convenience by $T_k$, decompose any vector x in G into the direct sum $x = \sum x_k$ of its orthogonal projections into the two-planes $B_k$ and define Tx by the formula $Tx = \sum T_k x_k$. If dim G $= 2\nu$, the number of independent parameters entering into this construction is $\nu(\nu -1)$ for the following reason. The choice of the vector a implies $(2\nu -1)$ parameters, since only the direction of a counts. But in the two-plane spanned by a and b, once constructed, any direction can replace that of a, hence the plane (ab) is determined by $2\nu -2$ parameters. Next, the vector c has to satisfy two orthogonality conditions, hence the number of parameters entering into the choice of the plane is $2\nu - 4$. The total sum is $(2\nu -2) + (2\nu - 4) + \ldots + 2 = \nu(\nu -1)$.

We return now to the isometries L and $L^{-1}$ treated above and, in particular, to their representation by T, as given by the formulae (4.2.6). If $\bar{\theta}$ denotes one particular solution of (4.2.1), then all solutions can be written

$$(4.2.7) \qquad \theta_r = \pm\ (\bar{\theta} + 2r\pi\ ), \quad r = 0, \pm 1, \pm 2, \ldots$$

Each of the two sequences of $\theta_r$, corresponding respectively to the signs plus and minus in front of the parenthesis, yields the same value of $\sin\theta$ and the same transformation T. All the values of $\theta_r$ can figure in the formulae (4.2,6) that is we have simultaneously

$$(4.2.8) \qquad L = e^{\theta_r T}, \qquad L^{-1} = e^{-\theta_r T}$$

for all $\theta_r$ .

In order to fix ideas one may choose the so called reduced solution of (4.2.1), that is the solution $\bar{\theta}$ satisfying $0 < \bar{\theta} < \pi$.

This concludes our discussion of subspaces of the type $E_Q$.

If a subspace $E_1$ actually occurs in the decomposition according to the respective schemes, given in 3.5 and 3.7, of the original space E, and if $E_1$ has the minimal polynomial (t-1), that is, if $L = 1$ and, consequently, $L^{-1} = 1$ in $E_1$, we clearly have $L - L^{-1} = 0$. As also the corresponding limiting value of $\sin \theta$ is to be taken equal to zero, the formula (4.2.3) degenerates entirely. On the other hand, the degenerate skew-symmetric transformation $S = 0$ gives a natural solution of the equation $e^S = 1$.

If a subspace $E_{-1}$ actually occurs in any one of the decompositions just quoted and has the minimal polynomial (t + 1), we have in $E_{-1}$, as before in $E_1$ , that $L - L^{-1} = 0$. The quotient (4.2.3) giving T for subspaces of the type $E_Q$ again becomes indeterminate. We recall, however, that the condition for L to be connected with the identity is that $E_{-1}$ should be of even dimension and moreover space-like in the Lorentz case. Under these conditions it is easy to find skew-symmetric transformations S such that $e^S = -1$ in $E_{-1}$. We have only to take an arbitrary skew-symmetric square root T of (-1) in $E_{-1}$ (their construction is explained earlier in this section), and set

$$S_r = (2r + 1) \pi T, \qquad r = 0, \pm 1, \pm 2, \dots .$$

4.3. Considerations similar to those at the beginning of section 4.2. apply to the hyperbolic rotation $L$ in the time-like two-plane $P$ treated in 3.9. From (3.9.6) we infer.

$$(4.3.1)\quad \begin{array}{ll} \frac{1}{2}(L + L^{-1})\, a = \mathrm{ch}\,\theta\,.a, & \frac{1}{2}(L + L^{-1})b = \mathrm{ch}\,\theta\,.b, \\ \frac{1}{2}(L - L^{-1})a = \mathrm{sh}\,\theta\,.b, & \frac{1}{2}(L - L^{-1})b = \mathrm{sh}\,\theta\,.a. \end{array}$$

If we define the linear transformation $T$ of $P$ into itself by

$$(4.3.2)\qquad T = \frac{L - L^{-1}}{2\mathrm{sh}\,\theta}$$

then $T$ is again skew-symmetric by (4.1.9) and, since by (4.3.1), $Ta = b$, $Tb = a$, we get the basic relation

$$(4.3.3)\qquad T^2 = 1,$$

which yields

$$(4.3.4)\qquad e^{\theta T} = \mathrm{ch}\,\theta + \mathrm{sh}\,\theta\,.T.$$

Formulae (3.9.6) and (4.3.4) give

$$L = \mathrm{ch}\,\theta + \mathrm{sh}\,\theta\,.T = e^{\theta T}, \quad L^{-1} = \mathrm{ch}\,\theta - \mathrm{sh}\,\theta\,.T = e^{-\theta T}.$$

An isometry $L$ of the whole Lorentz space $E$ in whose decomposition according to section 3.9 we have $\lambda < 0$, is not connected with the identity. Indeed, each half of the two generators invariant under $L$ is carried over into the other. From $(Ll \mid Lx) = (l \mid x)$ (or $(Lm \mid Lx) = (m \mid x)$ ) it follows that every positive time-like or light-like vector becomes negative and vice-versa, which implies that the two halves of the light-cone are interchanged. Clearly such a transformation is not connected with the identity.

4.4. We turn now to the parabolic case, treated in 3.10 and 3.11, consisting of such isometries of spaces $E_\varepsilon$ with $\varepsilon^2 = 1$, whose minimal polynomial is $(t - \varepsilon)^3$. For reasons analogous to those explained at the end of the previous section, we have to limit ourselves to the case $\varepsilon = 1$. So let $L$ be an isometry of the space $E_1$ with the minimal polynomial $(t - 1)^3$, which implies $(L - 1)^3 E_1 = 0$. We write this assumption in the form

(4.4.1) $$(L - 1)^3 = 0$$

(where by $L$ we mean the restriction of $L$ to $E_1$).

We shall prove that there exists a skew-symmetric transformation $T$ such that $L = e^T$, $L^{-1} = e^{-T}$. We first set

(4.4.2) $$A = \frac{L + L^{-1}}{2} .$$

Then $A - 1 = \frac{1}{2}L^{-1}(L^2 - 2L + 1) = \frac{1}{2}L^{-1}(L - 1)^2$.

Hence, on account of (4.4.1),

(4.4.3) $$(A - 1)^2 = 0.$$

Next we set

(4.4.4) $$T = \frac{L - L^{-1}}{2}$$

$T$ is obviously skew-symmetric. Moreover

(4.4.5) $$T^3 = 0, \qquad A = 1 + \frac{T^2}{2} .$$

The first relation follows, by virtue of (4.4.1), from $T = \frac{1}{2}(L - L^{-1}) = \frac{1}{2}(L^{-1} + 1)(L - 1)$, while the second follows from $A^2 - T^2 = 1$, and from (4.4.3).

Combining the second of the equations (4.4.5) and the equations $L = A + T$, $L^{-1} = A - T$, we get

(4.4.6) $L = 1 + T + \frac{1}{2}T^2, \quad L^{-1} = 1 - T + \frac{1}{2}T^2.$

On account of the first equation (4.4.5), this can be written

(4.4.7) $L = e^T, \quad L^{-1} = e^{-T}.$

Since $T$ defined by (4.4.4) is skew-symmetric, the transformations

$$L(s) = e^{sT} = 1 + sT + \frac{1}{2}s^2T^2$$

are isometries, as proved on p. 130. Moreover, $L(s+t) = L(s)L(t)$, that is the $L(s)$ form a one-parameter Abelian group.

Let now $x$ be any fixed vector in $E_1$. Then also $T_x$ and $T^2x$ are fixed, and

(4.4.8) $L(s)x = x + s.(Tx) + \frac{1}{2}s^2.(T^2x),$

which shows that, with $s$ varying from $-\infty$ to $+\infty$, $L(s)x$ describes a parabola. This explains at last the term parabolic rotation, used from time to time in the preceding. For a detailed description of the above parabolas see 4.7.

4.5 In the preceding we have constructed a skew-symmetric $T$ starting from a parabolic rotation $L$ and have derived the properties of $T$ from the known properties of $L$. Here we go the opposite way. We define a parabolic skew-symmetric $T$, investigate its properties and give in the next section a new proof of the fact that to every parabolic rotation $L$ there exists a $T$ such that $L = e^T$. We have two reasons for doing this. First, it gives us a better insight into the structure of parabolic rotations, studied relatively little up to now, secondly the skew-symmetric parabolic transformations are much easier to construct than the parabolic isometries. We start with the following theorem, the parallel of the theorem of 3.11.

*Let* $E$ *be a Lorentz space of dimension* $m \geq 3$, $\Pi$ *a light-like two-plane in* $E$, and $g = \operatorname{rad} \Pi$. *There exists one and, up to a* (non zero) *constant factor, only one skew-symmetric transformation* $T$ *such that* $TE = \Pi$. *Such a transformation* $T$ *has the further properties* $T\Pi = g$, $Tg = 0$. This implies in particular that the minimal polynomial of $E$ relative to $T$ is $t^3$, that of $\Pi$ is $t^2$ and that of $g$ is $t$.

The proof is analogous to that of the theorem of 3.11 but is easier. As there, we first analyze the implications of the assumption that there exists a transformation $T$ with the required properties. So let $T$ be skew-symmetric, and let $TE = \Pi$.

We first prove that $E' = \Pi^{\perp}$ is the subspace annihilated by $T$. We always have $(Tx \mid u) + (Tu \mid x) = 0$. Hence $(Tx \mid u) = 0$ and $(x \mid Tu) = 0$ imply each other. Let $Tu = 0$, then $(x \mid Tu) = 0$ and therefore $(u \mid Tx) = 0$ for an arbitrary $x \in E$. This means that $u$ is orthogonal to $\{Tx\} = \Pi$, thus $u \in \Pi^{\perp} = E'$. Conversely, let $u \in \Pi^{\perp}$. Then $(Tx \mid u) = 0$ and thus $(x \mid Tu) = 0$ for an arbitrary $x \in E$, which implies that $Tu = 0$. This proves that $TE' = T\Pi^{\perp} = 0$.

Since $g = E' \cap \Pi$, it follows that $Tg = 0$, and, more precisely, that if $y \in \Pi$, $Ty = 0$ if and only if $y \in g$.

Let $y \in \Pi$ but $y \notin g$. Then $u \in \Pi$ and $(y \mid u) = 0$ imply $u \in g$. Indeed, all vectors $u \in g$ satisfy the relation

$(y \mid u) = 0$, while $u = y$ does not satisfy it, because $(y \mid y) < 0$. Since $g$ and $y'$ span $\Pi$, our assertion follows.

We now prove that $T\Pi = g$. If $y \in g$, then $Ty = 0$, as proved above, hence in this case, $Ty \in g$. Next let $y \in \Pi$ but $y \notin g$. Then $Ty \in TE = \Pi$. On the other hand, from the skew-symmetry of $T$ it follows that $(y \mid Ty) = 0$. Hence, on account of what we just proved, $0 \neq Ty \in g$. Thus $0 \neq T\Pi \subset g$. From this and from $\dim g = 1$ follows that $T\Pi = g$. Hence $T^2E = T\Pi = g$ and $T^3E = T^2\Pi = Tg = 0$.

Summing up, we have as consequences of our assumptions

$$(4.5.1) \quad TE = \Pi \ ; \quad T^2E = T\Pi = g; \quad T^3E = T^2\Pi = Tg = 0.$$

This tells us that the respective minimal polynomials of $E, \Pi, g$ are $t^3$, $t^2$, $t$.

Since $TE' = 0$ and $T^2\Pi = 0$, also $T^2(E' + \Pi) = 0$. The sum $E' + \Pi = g$ (denoted by $T_g$ in section 3.6 ) is the $(m-1)$-dimensional tangent plane to the light cone along $g$. Since $T^2 g^{\perp} = 0$, $T^2E \neq 0$ and $\dim E = m$, plainly $g^{\perp}$ is exactly the subspace of $E$ annihilated by $T^2$.

Having examined the consequences of the assumption that $T$ is skew-symmetric and $TE = \Pi$, we now construct such a transformation $T$.

Let $y$ and $z$ be fixed vectors, $y \in \Pi$, $y \notin g$, $z \in g$, $z \neq 0$. There exists one and only one skew-symmetric $T$ such that $TE = \Pi$, $Ty = z$.

Proof. We first establish the existence. On account of our above findings, we set $Tv = 0$ for all $v \in E' = \Pi^{\perp}$ and in particular $Tz = 0$, for the fixed $z \in g \subset E'$. For the fixed

$y \in \Pi$, $y \notin g$ we set $Ty = z$. The $(m-1)$-dimensional tangent plane to the light cone along $g$ is spanned by $y$ and $E'$ (which is $(m-2)$-dimensional). We define $T$ by linearity, that is we set $T(\lambda y + v) = \lambda z$, for all real $\lambda$ and all $v \in E'$. Next, let $u \in E$ be a fixed vector outside the tangent plane, that is $(u \mid z) \neq 0$. We define $Tu$ by the equations

$(u \mid Tu) = 0$, $(y \mid Tu) + (Ty \mid u) = 0$, $(v \mid Tu) + (Tv \mid u) = 0$ or

(4.5.2) $(y \mid Tu) = -(z \mid u)$, $(u \mid Tu) = 0$, $(v \mid Tu) = 0$.

Since the $m$ vectors $u, y$ and $v_1, v_2, \dots, v_{m-2}$, where the $v_k$ form a basis of $E'$, are linearly independent, these equations have a unique solution $Tu$. Moreover $Tu \neq 0$, since $(y \mid Tu) = -(z \mid u) \neq 0$.

For an arbitrary vector $x \in E$ we write $x = \mu . u + \lambda . y + v$, $v \in E'$, and define $Tx$ by linearity, that is $Tx = \mu Tu + \lambda z$. We claim that $T$ is skew-symmetric. Indeed, by virtue of (4.5.2),

$$\begin{aligned}(x \mid Tx) &= (\mu u + \lambda y + v \mid \mu Tu + \lambda z)\\ &= \lambda \mu (u \mid z) + \lambda \mu (y \mid Tu) + \lambda^2 (y \mid z) + \lambda (v \mid z)\\ &= \lambda \mu (u \mid z) - \lambda \mu (u \mid z) + \lambda^2 (y \mid z) + \lambda (v \mid z)\\ &= \lambda^2 (y \mid z) + \lambda (v \mid z).\end{aligned}$$

However, $(y \mid z) = 0$ and $(v \mid z) = 0$, since $y \in \Pi$, $v \in E'$ and $z \in g = \operatorname{rad}(\Pi + E')$. Thus $(x \mid Tx) = 0$, that is $T$ is skew-symmetric, as asserted.

We still have to prove that $TE = \Pi$ or, in other words, that $Tu$ and $z$ span $\Pi$. Clearly, $0 \neq z \in g = \operatorname{rad} \Pi \subset \Pi$. On the other hand, $Tu \in \Pi = E'^{\perp}$, since $(Tu \mid v) = 0$ for an arbitrary $v \in E'$. Moreover

$Tu \notin g$, since $(y \mid Tu) = -(z \mid u) \neq 0$. Hence $z$ and $Tu$ are linearly independent and thus they span the two-plane $\Pi$.

Finally the uniqueness of the transformation $T$ follows from the uniqueness of $Tu$.

We have shown that there exists one and only one skew-symmetric transformation $T$ such that $TE = \Pi$ and $Ty = z$, where the given couple $y, z$ is such that $y \in \Pi$, $y \notin g$, $z \in g$, $z \neq 0$. We now show that if the requirement $Ty = z$ is dropped, $T$ is still determined up to a constant factor.

Let $\hat{T}$ be any skew-symmetric transformation such that $\hat{T}E = \Pi$. Then, as proved on p. 141, we have $0 \neq \hat{T}y \in g$ for any $y \in \Pi$, $y \notin g$. Thus keeping the above vector $y$, we find that $\hat{T}y = \hat{z} \in g$, $\hat{z} \neq 0$, and hence $\hat{z} = sz$, with a convenient real number $s \neq 0$, and our former $z$. The transformation $(s^{-1}\hat{T})$ is skew-symmetric, $(s^{-1}\hat{T})E = \Pi$, $(s^{-1}\hat{T})y = z$, hence by the above uniqueness theorem, $\hat{T} = sT$.

4.6. The results of the preceding section enable us to give a short proof of the fact established in section 4.4 that to every parabolic rotation $L$ there exists a skew-symmetric transformation $T$ such that $e^T = L$ and $TE = \Pi$. We shall also prove that $T$ is unique.

According to 3.10 or 3.11 we have $(L-1)E = \Pi$, $\Pi$ being a light-like two-plane. We again set rad $\Pi = g$ and take a fixed vector $y$ such that $y \in \Pi$, $y \notin g$. We have $(L-1)y = z$, with $z \in g$, $z \neq 0$. It follows that $(L-1)z = 0$. As proved in 3.11, the requirements that $L$ is an isometry, that $(L-1)E = \Pi$ and

$(L-1)y = z$ determine $L$ in a unique way.

According to 4.5 there exists one and only one skew-symmetric transformation $T$ such that $TE = \Pi$, $Ty = z$; it also follows that $Tz = 0$ and that $T^3 = 0$. We assert that $L = e^T$. Indeed, $e^T$ is an isometry, as proved in 4.1. Moreover, $e^T = 1 + T + \frac{1}{2}T^2$. We know that $TE = \Pi$, $T^2E = g = \operatorname{rad} \Pi$, where $g$ is a proper subspace of $\Pi$. Hence $(e^T-1)E = TE + \frac{1}{2}T^2E = \Pi$. Next, $Ty = z$, $T^2y = Tz = 0$, hence $(e^T-1)y = z$. From these arguments our assertion follows. The uniqueness follows from the fact that $(e^T-1)E = \Pi$ and $TE = \Pi$ are valid simultaneously, and the same is true for $(e^T-1)y = z$ and $Ty = z$.

4.7. We give here a detailed description of the parabolas considered in 4.4. There it was shown that the vector $x$ being fixed, and the parameter $s$ varying, the point (the position vector)

$$(4.7.1) \quad L(s)x = e^{sT}x = x + s(Tx) + \tfrac{1}{2}s^2(T^2x)$$

describes a parabola. Each parabola passes through the point $x$ from which it is derived and lies in a light-like two-plane parallel to $\Pi$. All the parabolas have the same point at infinity, identical to the point at infinity of the light-like line $g$. These assertions follow from $Tx \in \Pi$ and $T^2x \in g \subset \Pi$, and $Tx \notin g$, in general; all this was shown in 4.5. From the facts established there, it also follows that the parabola corresponding to $x$ degenerates into a straight line, parallel to $g$, if and only if $x$ lies in $g^\perp$ (that is the tangent hyperplane along $g$ to the light cone), and in particular if $x$ lies in $\Pi$, moreover that this parabola reduces to the point $x$ itself, if and only if $x$ lies in $E' = \Pi^{\perp}$, and in particular, if $x$ lies in g.

4.8. We once more recall the decomposition of a given vector space $E$ into the direct sum of mutually orthogonal sub-spaces, invariant under a given isometry $L$ as described in 3.5 and 3.7 respectively in the Euclidean case and in the Lorentz case. In the preceding we have obtained the representation of the restrictions of $L$ to these sub-spaces. It remains to obtain the representation by infinitesimal transformations of the isometry $L$ for the entire space.

For each restriction of $L$ we have found one or several skew-symmetric transformation $S$ such that in the corresponding subspace the formulae

$$(4.8.1) \qquad L = e^{S}, \quad L^{-1} = e^{-S}$$

were valid. In a schematical way, denote the subspaces by $E_A, E_B, E_C, \ldots,$ the corresponding restricted isometries and skew-transformations by $L_A, L_B, L_C, \ldots$ and $T_A, T_B, T_C, \ldots$ respectively. Define the linear transformation $S$ of $E$ into itself by setting $S = S_M$ in $E_M$ for $M = A, B, C, \ldots.$ Then clearly (4.8.1) holds in the entire space $E$.

We claim that $S$ is a skew-symmetric transformation in $E$. Write any vector $x \in E$ as the direct sum $x = \sum x_M$ with $x_M \in E_M$. Then $Tx = T(\sum x_N) = \sum Tx_N = \sum T_N x_N$. Hence $(x \mid Tx) = \sum \sum (x_M \mid T_N x_N)$. Here every term vanishes. Indeed if $M = N$, then $(x_M \mid T_M x_M) = 0$ on account of the skew-symmetry of $T_M$. If $M \neq N$, $x_M \perp T_N x_N$, since $x_M \in E_M$, $T_N x_N \in E_N$ and $E_M \perp E_N$. Herewith the proof is complete.

We add a few comments on the multivalence of the infinitesimal transformation S generating L in the sense of formulae (4.8.1). To this effect we have to examine the infinitesimal transformations generating the restrictions of L to the different types of subspaces.

We begin with a subspace of the type treated in 4.2 with the minimal polynomial $Q(t) = t^2 + \beta t + 1 = t^2 - 2\cos\theta + 1$. We have found the infinitesimal transformations

$$(4.8.2) \qquad S_r = \theta_r \frac{L - L^{-1}}{2\sin\theta_r}, \quad \theta_r = \pm(\bar{\theta} + 2r\pi),$$

$$r = 0, \pm 1, \pm 2, \ldots$$

where $\bar{\theta}$ is an arbitrary fixed solution of the equation $\cos\theta = -\frac{1}{2}\beta$. Since $\theta_r/\sin\theta_r$ has the same value for $\pm\theta_r$, the $S_r$ form a unique sequence with the period $\pi(L - L^{-1})/\sin\bar{\theta}$.

In the hyperbolic case treated in 4.3, the skew-symmetric transformation S is entirely determined by the condition $e^S = L$. The explicit solution reads on account of (4.3.2)

$$(4.8.3) \qquad S = \theta T = \frac{\theta(L - L^{-1})}{2\,\mathrm{sh}\,\theta}$$

The hyperbolic angle was determined in 3.9 by the equation $e^\theta = \lambda$, where $\lambda$ was one of the two eigenvalues $\lambda$ and $\lambda^{-1}$ belonging to the plane P. By this $\theta$ is entirely determined. However, the choice $e^\theta = \lambda^{-1}$, replacing $\theta$ by $-\theta$, would have been equally possible. This alteration of

$\theta$ does not change $\theta / \mathrm{sh}\, \theta$.

In the Euclidean case a subspace $E_{+1}$ has the minimal polynomial $t - 1$, while in the Lorentz case the minimal polynomial is either $t - 1$ or $(t - 1)^3$. In the case $t - 1$ the restriction of $L$ to $E_{+1}$ is the identity and we could assign $S = 0$ as the natural solution of the equation $e^S = 1$. In the case $(t - 1)^3$, according to 4.4 - 4.6, the equation $e^s = L$ admits the solution $S = \frac{1}{2}(L - L^{-1})$.

While the indeterminacy relative to the subspaces $E_Q$ was of ordinary logarithmic character, implying certain periods, an indeterminacy of different character occurs in a subspace $E_{-1}$, where one has to consider arbitrary skew-symmetric solutions of the equation $e^S = -1$. We have found the solutions $S_r = (2r+1)\pi T$ where $T$ was an arbitrary skew-symmetric square root of $(-1)$ in $E_{-1}$. Now the subspace $E_{-1}$ is always determined by $L$, and so also is $\{S_r\}$ if $\dim E_{-1} = 2$. But if $\dim E_{-1} > 2$, $L$ in no way determines the choice of $T$ (see 4.2) and the corresponding $\{S_r\}$.

It is well known that a necessary and sufficient condition for Cayley's parameterization of isometries is that the isometry in question should not have the eigenvalue (-1).

In contrast to the Cayley formulae, the eigenvalue (-1) does not prevent the representation of $L$ by an infinitesimal transformation, but in exchange it produces a high degree of indeterminacy, if $\dim E_{-1} > 2$.

The various infinitesimal transformations $S$ figuring in (4.8.1) can, to a certain extent, be expressed in an explicit way as polynomials of the given isometry $L$. The requisite tools are furnished by the "interpolation formula" (3.2.12) and the remarks following that formula. Indeed, the restrictions $S_M$ of $S$ to any of the subspaces schematically denote by $E_M$ are scalar multiples of the corresponding $L_M - L_M^{-1}$, except in the subspace $E_M = E_{-1}$, where $S_M \neq 0$, while $L_M - L_M^{-1} = 0$. Thus, for $M \neq -1$, the transformation $S_M$ is a certain polynomial $h^M$ of $L_M$. Hence, according to the formula and the remarks just quoted, there exists a polynomial $h(t)$ (of degree less than than of the minimal polynomial $m(t)$ of the space $E$ relative to $L$) such that $h(L) = S$ in $E_{-1}^{\perp}$ (= the direct sum of all subspaces $E_M$ except $E_{-1}$), and $h(L) = 0$ in $E_{-1}$. Finally, in the entire space $E$ we have $S = h(L) + S'$, where $S' = (2r + 1)\pi T$ in $E_{-1}$ and $0$ in $E_{-1}^{\perp}$, and $T$ is a skew-symmetric square root of $(-1)$ in $E_{-1}$. Obviously the whole situation becomes simpler if no subspace $E_{-1}$ occurs.

We add the following remark, restricting ourselves to subspaces of type $E_Q$. The transformations $\pm \theta_r T$ figuring in the exponents in formulae (4.2.8) are scalar multiples of a certain square root $T$ of $(-1)$ in $E_Q$. If dim $E_Q > 4$, we can decompose $E_Q$ into two-planes $B_k$, invariant relative to $L$ (see p. 135 ). Each of the skew-symmetric square roots of $(-1)$ corresponding to these two-planes can be treated separately and gives rise to a sequence $\pm \theta_r T_k$, where the $\theta_r$ can, for each $k$, be chosen independently. Denoting them by $\theta_{rk}$ and writing $S = \theta_{rk} T_k$ in $B_k$, one still has $L = e^S$, $L^{-1} = e^{-S}$. However, if the $\theta_{rk}$ are not identical for different k, the transformation S just constructed can not be expressed as a polynomial of L.

We conclude this section by a statement about the minimal polynomials of skew-symmetric transformations, which follows readily from the connection between this kind of transformations and isometries. The direct proofs, which we recommend to the reader as exercises, are analogous to, but considerably simpler than, the corresponding proofs for isometries in sections 3.5, 3.7-3.10.

The minimal polynomial m(t) of a skew-symmetric transformation in a space E of dimension n is of the following form.

a) Euclidean case:

$$m(t) = t^{p} \prod_{k=1}^{h} (t^2 + \beta_k)$$

where $p = 0,1$; $h = 0,1,\dots$; the $\beta_k$ are distinct positive numbers; $p + 2h \leqslant n$;

b) Lorentz case:

$$m(t) = t^{p} (t^2 - \gamma)^{r} \prod_{k=1}^{h} (t^2 + \beta_k),$$

where $p = 0,1,3$; $r = 0,1$; $h = 0,1,\dots$; $\gamma$ is a positive number; the $\beta_k$ are distinct positive numbers; $p + 2r + 2h \leqslant n$; p and r cannot vanish simultaneously; the combination $p = 3$, $r = 1$ cannot occur.

Both cases: The invariant subspaces of E, which have the various relatively prime factors as minimal polynomials, are mutually orthogonal.

4.9. Sylvester's law of inertia asserts that when a real quadratic form is transformed by any non-singular real linear transformation into a sum of squares with coefficients ± 1, then the number $k$ of the positive terms and the number $l$ of the negative terms are independent of the transformation. With a slight change of the usual terminology we say that the couple $(k, l)$ of these two numbers constitute the signature of the form. From our present point of view, we do not even distinguish between the signatures $(k, l)$ and $(l, k)$. Thus, a sum of squares with all signs equal will be said to have Euclidean signature, and a sum with all signs equal but one will be said to have Lorentz signature. The same denominations will be used for the corresponding spaces. In this terminology we can say that any isometry connected with the identity of a space with Euclidean or Lorentz signature can be generated by a single infinitesimal transformation. We shall show that this is not generally true for any other signature. We only sketch the proof which requires a certain familiarity with the theory of elementary divisors.

Consider the metric form

$$(x|x) = x_1x_4 + x_2x_3 + x_3x_2 + x_4x_1 \tag{4.9.1}$$

and the corresponding scalar product

$$(x|y) = x_1y_4 + x_2y_3 + x_3y_2 + x_4y_1 . \tag{4.9.2}$$

Consider also the matrix of order four

$$T = \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix} , \qquad \text{where} \qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} . \tag{4.9.3}$$

As easily seen $T$ is skew-symmetric relative to the above metric, that is $(x|Tx) = 0$. Consequently, as proved in 4.1, the transformations $L(s) = e^{sT}$ generated by $T$ are isometries of the metric; as usual $s$ is a real parameter. Clearly

$$A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},$$

and from this

$$(4.9.4) \qquad L(s) = e^{sT} = \begin{pmatrix} e^{s}A(s) & 0 \\ 0 & e^{-s}A(-s) \end{pmatrix}, \text{ where } A(s) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

The isometries $L(s)$ are connected with the identity since they are generated by an infinitesimal transformation. This also appears from the explicit form (4.9.4) of $L(s)$.

Like every quadratic metric, the above particular one admits the isometry $x \rightarrow (-x)$. The remarkable fact is, however, that here, in contrast to the Lorentz case, this isometry is connected with the identity, as will turn out in a moment.

The above metric is the so-called ultra-hyperbolic metric in four dimensions. We introduce a new basis by setting.

$$x_1 = z_1 + z_4, \quad x_4 = z_1 - z_4, \quad x_2 = z_2 + z_3, \quad x_3 = z_2 - z_3.$$

We then have

$$(4.9.5) \qquad (x|x) = 2(z_1^2 + z_2^2 - z_3^2 - z_4^2) = 2B(z,z).$$

Now, clearly $x \rightarrow (-x)$ is the same transformation as $z \rightarrow (-z)$. Our assertion concerning the first transformation will be proved if we can link the second transformation to the identity by a continuous passage preserving the form $B(z,z)$. But this is an easy task, as seen from the formulae

$$z_1' = \cos\theta . z_1 + \sin\theta . z_2, \qquad z_3' = \cos\theta . z_3 + \sin\theta . z_4,$$

$$z_2' = -\sin\theta . z_1 + \cos\theta . z_2, \qquad z_4' = -\sin\theta . z_3 + \cos\theta . z_4,$$

$\theta$ varying from $0$ to $\pi$ .

Recalling 4.2, we see that the rotations above are generated by an infinitesimal transformation whose effect on the couples $(z_1, z_2)$ and $(z_3, z_4)$ is represented by the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Denoting the infinitesimal transformation acting on the whole space by $U$, we have $-x = e^{\pi U} x$.

From the preceding it is clear that the isometries $-L(s)$ are also connected with the identity; moreover they can be generated by two infinitesimal transformations $T$ and $U$,

$$-L(s) = e^{\pi U} e^{sT} = e^{sT} e^{\pi U} .$$

This being the case, we now show that none of these transformations $-L(s)$, with $s \neq 0$, can be generated by a single infinitesimal transformation. More precisely, we show that, $s$ being fixed, the weaker assumption $-L(s) = e^{V}$ for some fixed $s \neq 0$ with a real matrix $V$ leads to a contradiction.

The assumption $-L(s) = e^{V}$ implies that $-L(s) = M^2$, where the matrix $M = e^{\frac{1}{2}V}$ is real. We show that even this still weaker assumption leads to a contradiction. The eigenvalues of $-L(s)$ are clearly $-e^{s}$ and $-e^{-s}$, which are different, since $s \neq 0$. Hence the possible eigenvalues of $M$ are $\pm ie^{\frac{1}{2}s}$ and $\pm ie^{-\frac{1}{2}s}$, each of these eigenvalues occurring either with one of the signs $\pm$ or both. However, only the latter case is possible, since non real eigenvalues of a real matrix must occur in conjugate complex pairs. Hence $M$ should have four distinct eigenvalues, which would imply that all its elementary divisors are simple. This contradicts the fact that $-L(s) = M^2$ has two elementary divisors of order 2.

It is apparent that our counter-example was based on the existence of an infinitesimal transformation having two distinct eigenvalues with multiple elementary divisors and also on the topological fact that reflection in the origin is connected with the identity in the group of isometries.

What precedes is particularly interesting in the light of the problem posed by von Neumann (Math. Zeitschr. 30 (1929), p. 26) and relative to an arbitrary group G of linear transformations, whether it is possible to generate any particular element of G connected with the identity by at most two infinitesimal transformations.

In spaces of dimension $\leqslant 3$ no similar counter-example can be constructed owing to the fact that such a space is always of Euclidean or Lorentz signature. On the other hand, the above counter-example relative to four dimensions can serve to prove the same negative result in a space of arbitrary dimension $> 4$. Indeed, in such a space any sum of squares, which is not of Euclidean or Lorentz type, includes at least two positive and two negative terms. This being the case, the construction of a counterexample can be carried out in the following way. The variables corresponding to the four terms mentioned are transformed according to the above isometry -L(s). The remaining variables are subjected to an isometry connected with the identity and having distinct eigenvalues which are also different from the eigenvalues $-e^{s}$ and $-e^{-s}$ of -L(s). Then it follows as before that the matrix corresponding to the total isometry cannot be the square of a real matrix. With this the proof is complete.

Spaces regular with respect to some metric which is of neither Euclidean nor Lorentz signature, can in many ways be characterized by intrinsic properties. One such property is that they contain some totally isotropic subspace of dimensions $\geqslant 2$.

4.10. For the sake of completeness and for the convenience of the reader, in particular if he happens to be a physicist, we give here the representation of skew-symmetric transformations by *matrices*. This will also facilitate an important identification needed in the next section.

Le $E$ be a real vector space of $n$ dimensions, regular with respect to a given metric $(x|y)$. Let $\{e_k\}$ be a basis of $E$ and $\{e^k\}$ the complementary basis constructed on p. 94. Any vector can be expressed by either of these two bases; with the usual summation convention, we write for a generic vector $x$

(4.10.1) $$x = x^k e_k = x_k e^k, \text{ where } x^k = (x|e^k) \text{ and } x_k = (x|e_k).$$

Let the *transformation* $T$ of $E$ into itself be *skew-symmetric*, that is $(x|Tx) = 0$ or equivalently $(x|Ty) + (y\ Tx) = 0$. Set

(4.10.2) $$Te_k = \alpha_{rk} e^r \quad \text{and} \quad Te^k = \alpha^{rk} e_r .$$

Here and in the sequel the raising and the lowering of indices is to be performed according to the ordinary rules of tensor calculus by means of the metric matrices $(g_{jk})$ and $(g^{jk})$ with

$g_{jk} = (e_j \mid e_k)$ and $g^{jk} = (e^j \mid e^k)$, where the two matrices are inverse to each other.

The relations $(e_j \mid Te_k) + (e_k \mid Te_j) = 0$ and $(e^j \mid Te^k) + (e^k \mid Te^j) = 0$ yield respectively

$$(4.10.3) \qquad \alpha_{jk} + \alpha_{kj} = 0 \quad \text{and} \quad \alpha^{jk} + \alpha^{kj} = 0,$$

that is the matrices $(\alpha_{jk})$ and $(\alpha^{jk})$ are skew-symmetric. It is readily seen that the converse is true, namely that the skew-symmetry of either of the above matrices is a sufficient condition for the skew-symmetry of the transformation $Tx$. We omit the proof.

The transformation $Tx$ can be expressed by the coefficients $\alpha_{jk}$ or $\alpha^{jk}$ as follows:

$$(4.10.4) \qquad Tx = x^k Te_k = x^k \alpha_{rk} e^r = (\alpha_{rk} x^k) e^r,$$

or

$$(4.10.5) \qquad Tx = x_k Te^k = x_k \alpha^{rk} e_r = (\alpha^{rk} x_k) e_r .$$

The coordinates $x^k$ are transformed according to the following rule

$$(Tx)^k = (Tx \mid e^k) = -(x \mid Te^k) = -(x \mid \alpha^{rk} e_r) = -\alpha^{rk} x_r = \alpha^{kr} x_r .$$

This formula with the analogous formula for $(Tx)_k$ and the notations $(Tx)^k = Tx^k$, $(Tx)_k = Tx_k$, enable us to write

$$(4.10.6) \qquad Tx^k = \alpha^{kr} x_r , \qquad Tx_k = \alpha_{kr} x^r .$$

These formulae are in perfect agreement with (4.10.4) and (4.10.5) and could have been derived from those.

Finally, we write, in notations also familiar from tensor calculus,

$$(4.10.7)\qquad Te_k = \alpha^{r\cdot}_{\cdot k}\, e_r, \qquad Te^k = \alpha^{k\cdot}_{r\cdot}\, e^r,$$

$$Tx^k = \alpha^{k\cdot}_{\cdot r}\, x^r, \qquad Tx_k = \alpha^{\cdot r}_{k\cdot}\, x_r .$$

Obviously the corresponding mixed matrices are not, in general, skew-symmetric any more.

4.11. After a long intermission we now return to our Clifford algebras, concentrating upon the bivectors of such an algebra in their role as skew-symmetric operators. We recall that a Clifford algebra is an associative algebra based on either of the equivalent relations $x^2 = (x \mid x)$, $xy + yx = 2(x \mid y)$, for arbitrary vectors. Here we consider the Clifford algebra over the real field of scalars. In Chapter VI the complex Clifford algebra will play an essential role. The vector space $E$ is the same as in the preceding section. For the moment all vectors in $E$, with their products and repeated products, are considered as Clifford numbers.

We deduce here a number of relations between certain products. Let $a, b, c$ be vectors; we show that $abc - bca$ is a vector and $abc - cba$ is a trivector. We have

$$(4.11.1)\qquad abc - bca = (ab + ba)c - b(ac + ca)$$
$$2(a \mid b)c - 2(a \mid c)b,$$

which proves our first assertion. Next we recall that the commutator of any $p$ quantities $U_1, U_2, \ldots, U_p$ in the Clifford algebra is defined by

(4.11.2) $$[U_1, U_2, \dots, U_p] = \varepsilon^{i_1 i_2 \dots i_p} U_{i_1} U_{i_2}, \dots U_{i_p},$$

Where $\varepsilon^{i_1 i_2 \dots i_p}$ is the totally alternate tensor (cf.(1.1o.1)).

We write in particular

(4.11.3) $$[U, V] = UV - VU.$$

We shall now show that

(4.11.4) $$abc - cba = bca - acb = cab - bac = (1/3)[a, b, c].$$

The three differences are obtained from each other by cyclical permutations, their sum is equal to $[a, b, c]$ and so everything will be clear, if they are proved to be equal. We have, for example,

$$abc - cba - (bca - acb) = a(bc + cb) - (cb + bc)a =$$

$$= 2(b \mid c)a - 2(b \mid c)a = 0.$$

Now let $x$ be a vector, $F$ a bivector

(4.11.5) $$x = x^k e_k = x_k e^k, \quad F = \tfrac{1}{2}F^{rs} e_r e_s = \tfrac{1}{2}F_{rs} e^r e^s,$$

$$F^{rs} = -F^{sr}, \quad F_{rs} = -F_{sr}.$$

We also need the following form, which results from skew-symmetry,

(4.11.6) $$F = \tfrac{1}{4}F_{rs}(e^r e^s - e^s e^r).$$

On account of (4.11.5) and (4.11.1) we have

(4.11.7) $$Fx - xF = \tfrac{1}{2}F^{rs}(e_r e_s x - x e_r e_s) =$$

$$= F^{rs}\{(x \mid e_s)e_r - (x \mid e_r)e_s\} = F^{rs}x_s e_r - F^{rs}x_r e_s =$$

$$= 2F^{rs}x_s e_r = 2F_{rs}x^s e^r.$$

Next, using (4.11.6) and (4.11.4),

$$(4.11.8) \qquad Fx + xF = \tfrac{1}{4}F_{rs} \left\{ (e^r e^s x - x e^s e^r) - (e^s e^r x - x e^r e^s) \right\}$$

$$= (1/6) F_{rs} \left[ e^r, e^s, x \right] = (1/6) F_{rs} x_k \left[ e^r, e^s, e^k \right] =$$

$$= (1/6) \sum_{k,r,s} (F_{rs} x_k + F_{sk} x_r + F_{kr} x_s) \left[ e^r, e^s, e^k \right],$$

where the summation is extended to all combinations $(k,r,s)$ such that $k < r < s$.

In view of an application in section 4.17 we also write out $xF$ and $Fx$ explicitly. From (4.11.7) and (4.11.8) it follows that

$$(4.11.9) \qquad \left. \begin{matrix} xF \\ Fx \end{matrix} \right\} = \mp x_s F^{rs} e_r + \frac{1}{12} \sum_{k,r,s} (x_k F_{rs} + x_r F_{sk} + x_s F_{kr}) \left[ e^k, e^r, e^s \right].$$

We see that $xF$ and $Fx$ are *sums of a vector and a trivector*, whose explicit expressions are displayed by formula (4.11.9).

We now turn our attention to *the connection between skew-symmetric transformations and bivectors*. We set

$$(4.11.10) \qquad Sx = Fx - xF = \left[ F, x \right].$$

From the explicit formula (4.11.7) we obtain in particular

$$(4.11.11) \qquad Se_k = 2F_{rk} e^r, \qquad Se^k = 2F^{rk} e_r.$$

The skew-symmetry of $S$ follows from

$$(4.11.12) \qquad (x \mid Sx) = \tfrac{1}{2}(x.Sx + Sx.x)$$

$$= \tfrac{1}{2} \left\{ x(Fx - xF) + (Fx - xF)x \right\} = \tfrac{1}{2} \left\{ -x^2F + Fx^2 \right\} = 0,$$

since $x^2$ is a scalar. It also follows from the explicit formula (4.11.7), which gives $(x \mid Sx) = 2F^{rs} x_r x_s = 0$, since $F^{rs} = -F^{sr}$.

We shall show in section 4.18 that, with a convenient norm for Clifford numbers, the power series

$$e^{sF} = \sum_{n=0}^{\infty} \frac{s^n}{n!} F^n$$

is absolutely convergent and that its sum is a Clifford number possessing the ordinary properties of an exponential, in the same way as the exponential $e^{sT}$ considered in section 4.1.

We define

$$(4.11.13) \qquad L(s)x = e^{sF} x e^{-sF}$$

and prove that the transformations $L(s)$ form a one-parameter group of isometries generated by the infinitesimal transformation $Sx = Fx - xF$. We find, since $F$ and $e^{\pm sF}$ commute,

$$\frac{d}{ds} L(s)x = e^{sF} Fxe^{-sF} - e^{sF} xFe^{-sF} = e^{sF}(Fx - xF)e^{-sF}$$

$$= e^{sF} Sxe^{-sF} .$$

Hence by iteration

$$\frac{d^n}{ds^n} L(s)x = e^{sF} S^n xe^{-sF} ,$$

and in particular

$$\frac{d^n}{ds^n} L(s)x_{(s=o)} = S^n x$$

from which we conclude, as on p. 131 that

$$(4.11.14) \qquad L(s)x = e^{sS}x.$$

Our assertion then follows, by reasons given on p. 130 , from the skew-symmetry of $S$.

Summing up, we have proved that any bivector $F$ generates through (4.11.10) a skew-symmetric transformation and by (4.11.13) a one parameter group of isometries. The basic fact that conversely *to every skew-symmetric transformation* $T$ *there corresponds a bivector* $F$ *such that* $Tx = Fx - xF$, follows from formulae (4.10.2) and (4.11.11) if we set

$$\alpha_{ik} = 2F_{ik}. \tag{4.11.15}$$

With this choice we also have the identity

$$e^{sT}x = e^{sF}xe^{-sF}. \tag{4.11.16}$$

4.12. The results obtained in the preceding sections enable us to establish the following theorem which was one of our principal aims through Chapters III and IV.

*Let* $E$ *be a real vector space of Euclidean or Lorentz signature and* $C$ *the corresponding Clifford algebra. To every isometry* $L$ *of* $E$, *connected with the identity, there exists a bivector* $F$ *in* $C$ *such that*

$$Lx = e^{F}xe^{-F} \tag{4.12.1}$$

*for any vector* $x$. *This property does not belong to any other signature.*

The right hand side of (4.12.1) is to be interpreted in the Clifford sense. The theorem follows immediately from the fact that, as proved in 4.8, $L$ can be written as an exponential $e^{S}$ of a skew-symmetric $S$. According to the first formula (4.10.2), this transformation $S$ can be expressed by relations of the form $Se_k = \alpha_{rk}e^r$. Then the bivector $F = \frac{1}{2}F_{ik}e^ie^k$, related to $S$ by the formula

$$Sx = Fx - xF, \tag{4.12.2}$$

is on account of the formula (4.11.15) given by the relations $F_{ik} = \frac{1}{2} \alpha_{ik}$.

This proves the positive part or our assertion. The negative part follows from the corresponding negative result concerning skew-symmetric transformations given in section 4.9.

Returning to the positive part of our result, we know from section 4.8 that the skew-symmetric $S$ corresponding to a given isometry $L$ is not unique in general. To different skew-symmetric transformations $S$, there correspond different bivectors $F$. While all these $S$ give the same exponential $e^S$, the exponentials $e^F$ are only determined up to a factor $\pm 1$, as we shall now prove.

We first prove: If for two bivectors $F'$ and $F''$ we have identically in $x$,

$$e^{F'} x e^{-F'} = e^{F''} x e^{-F''} ,$$

then $e^{F''} = \pm e^{F}$. Set $e^{-F''} e^{F'} = \Lambda$. Then the above relation reads $\Lambda x = x \Lambda$, that is $\Lambda$ commutes with any vector $x$ and consequently with any element of the Clifford algebra $C$. Thus $\Lambda$ belongs to the center of $C$. If $\dim E$ is even, it follows immediately that $\Lambda$ is scalar (cf. 1.5). This follows also in the case when $\dim E$ is odd. Indeed, in this case the center of $C$ consists of the elements $\alpha + \beta P$ where $\alpha$ and $\beta$ are arbitrary scalars and $P$ is the pseudoscalar, that is a quantity of odd dimension (cf. 1.5). Now $\Lambda$ is obviously of even dimension, hence its belonging to the center of $C$ implies that it is a scalar.

It remains to show that $\Lambda = \pm 1$. We recall the antiautomorphism $c \rightarrow \tilde{c}$ of a Clifford algebra, termed reversion (cf p. 13), which, among other things, keeps the scalars unchanged and changes the signs of the bivectors. Hence, on one hand, $\tilde{\Lambda} = \Lambda$, and on the other hand $\tilde{\Lambda} = e^{\tilde{F}'} e^{-\tilde{F}''} = e^{-F'} e^{F''} = \Lambda^{-1}$. Hence $\Lambda = \Lambda^{-1}$ and thus $\Lambda = \pm 1$.

This indeterminacy is an intrinsic property of the representation of isometries by inner automorphisms like that given by formula (4.12.1). We return to this question in section 4.15 (see also sec. 2.4).

Finally we point out that the same number of parameters enters into an isometry and into a bivector (cf. 2.3). Indeed, an isometry of an n-dimensional vector space can be considered as an automorphism of the metric form $(x \mid x) = g_{ik} x^i x^k$. A linear transformation in $n$ variable includes $n^2$ coefficients; the metric form, because of its symmetry, includes $\frac{1}{2}n(n+1)$ coefficients, hence an isometry depends on $n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$ parameters. The same number of parameters enters into a bivector, because of its skew-symmetry. Hence the formula (4.12.1) displays the right number of parameters.

4.13. In the preceding section we used a basis to construct in a global way the bivector $F$, as a solution of the equation $Tx = Fx - xF$, where $T$ was a given skew-symmetric transformation. No basis whatever will be used in the following, where we shall describe another construction in which $F$ is built up from simple bivectors relative to two-planes invariant under $L$. We merely use and further develop the procedure given earlier (sec. 2.5) in connection with Hamilton's reflection principle. We recall our results and we make some slight changes in order to reach a complete agreement with the preceding sections. Also in what follows, only spaces of either Euclidean or Lorentz signature will be considered. More precisely, the signatures should be those following from the respective abstract definitions in 3.5 and 3.6, see also the metric forms (3.3.3) and (3.3.4).

We begin with a Euclidean two-plane $\Pi$ (cf. 2.5, 3.5, 4.2, 4.8). Let $a$ and $b$ be mutually orthogonal unit vectors, that is $(a \mid a) = (b \mid b) = 1$, $(a \mid b) = 0$, or in Clifford language $a^2 = b^2 = 1$, $ab + ba = 0$, hence $a^{-1} = a$ and $\frac{1}{2}(ab - ba) = ab$. We introduce the unit bivector $I = ba^{-1}$; here it is the same as $I = ba$, the advantage of the definition chosen will appear later. $I$ has the basic properties $I^2 = -1$, $Ia = -aI = b$, $Ib = -bI = -a$. The bivector $I$ anticommutes with every vector $x \in \Pi$, that is $Ix + xI = 0$. Up to a factor $\pm 1$ there is only one unit bivector in the plane; if $a'$ and $b'$ are another couple of mutually orthogonal unit vectors in $\Pi$, then $a' = \cos\varphi . a + \sin\varphi . b$, $b' = \pm(-\sin\varphi . a + \cos\varphi . b)$ and thus $ba = \pm b'a'$. We have

$$(4.13.1) \qquad e^{\theta I} = \cos\theta + I \sin\theta,$$

and

$$(4.13.2) \qquad e^{\theta I}a = \cos\theta . a + \sin\theta . b,$$
$$e^{\theta I}b = -\sin\theta . a + \cos\theta . b.$$

The effect of the rotation with angle $\theta$ on an arbitrary vector

$x \in \Pi$ is expressed by the formula

$$(4.13.3) \qquad x' = \mathrm{rot}_\theta\, x = e^{\theta I} x .$$

Since $x$ anticommutes with I, we have $e^{\frac{1}{2}\theta I} x = x e^{-\frac{1}{2}\theta I}$, hence the fundamental formula

$$(4.13.4) \qquad \mathrm{rot}_\theta\, x = e^{\frac{1}{2}\theta I} x e^{-\frac{1}{2}\theta I} .$$

This formula (but not (4.13.3) can be extended to an arbitrary vector $x$ in space. Writing $x = x_{\parallel} + x_{\perp}$, where $x_{\parallel}$ is the orthogonal projection of $x$ into $\Pi$ and $x_{\perp}$ is orthogonal to $\Pi$, $\mathrm{rot}_\theta\, x$, relative to $\Pi$, is in a natural way, defined by

$$(4.13.5) \qquad \mathrm{rot}_\theta\, x = \mathrm{rot}_\theta\, x_{\parallel} + x_{\perp} .$$

Since $x$ commutes with I, the formula (4.13.4) remains valid in this extended meaning. The passage $\theta \longrightarrow \theta + 2r\pi$ changes neither the rotation nor the exponential in (4.13.3), but it does change the exponentials in (4.13.4) by a factor $(-1)^r$, in perfect agreement with what has been seen before.

Let now $\Pi$ be a space-like two-plane in a Lorentz-space (cf. 3.8, 4.2, 4.8). Let a and b again be two mutually orthogonal unit vectors in $\Pi$. Here this means $(a \mid a) = (b \mid b) = -1$, $(a \mid b) = 0$, or $a^2 = b^2 = -1$, $ab + ba = 0$. We again define $I = ba^{-1}$ and have in the present case $I = -a^{-1}b = ab$ and again $I^2 = -1$, $Ia = b$, $Ib = -a$, and all the apparatus becomes identical to that developed in the Euclidean case.

Next let $\Pi$ be a time-like two-plane (cf. 3.9, 4.3). Taking here two mutually orthogonal unit vectors, one is time-like, the other space-like. We consequently have $a^2 = \pm 1$,

$b^2 = \mp 1$, $ab + ba = 0$. We set $I = ba^{-1}$. Then $Ia = b$, $Ib = ba^{-1}b = -b^2a^{-1} = -b^2a^{-2}a = -b^2a^2a$, hence $Ib = a$. Moreover $I^2 = ba^{-1}ba^{-1} = -b^2a^{-2}$, that is $I^2 = 1$. The bivector $I$ is unique, except for sign. If $a' = \mathrm{ch}\,\varphi.a + \mathrm{sh}\,\varphi.a$, $b' = \pm(\mathrm{sh}\,\varphi.a + \mathrm{ch}\,\varphi.b)$, we have $I' = \pm I$. The bivector $I$ anticommutes with every vector $x$ in $\Pi$ and commutes with every vector orthogonal to $\Pi$. The effect of the rotation in $\Pi$ with the hyperbolic angle $\theta$ upon any vector $x$ in the space can be written

$$\mathrm{rot}_\theta\, x = \mathrm{rot}_\theta\, x_\parallel + x_\perp = e^{\frac{1}{2}\theta I}\, x e^{-\frac{1}{2}\theta I},$$

with $e^{\pm\frac{1}{2}\theta I} = \mathrm{ch}\frac{1}{2}\theta \pm \mathrm{sh}\frac{1}{2}\theta .I$.

Finally, we consider a parabolic rotation related to a light-like-two-plane $\Pi$ (cf.3.7, e.10, 3.11, 4.4 - 4.7). The plane contains a light-like line $g = \mathrm{rad}\,\Pi$ (see loc.cit.). Let $a \in \Pi$, $a \notin g$, $b \in g$, $b \neq 0$. Then $(b\,|\,b) = 0$, $(a\,|\,a) < 0$, $(a\,|\,b) = 0$, or $b^2 = 0$, $a^2 < 0$, $ab + ba = 0$. We form the bivector $I = ba^{-1}$ (with $a^{-1} = a/(a\,|\,a)$). Then we have $I^2 = ba^{-1}ba^{-1} = -b^2a^{-2} = 0$. Hence the bivector $I$ cannot be normalized. We show that it is unique except for a scalar factor. Indeed with $a' = \lambda a + \mu b$, $b' = \alpha b$, we have $I' = \alpha\lambda^{-1}I$. Again the bivector $I$ anticommutes with every vector in $\Pi$. We have $Ia = b$, $Ib = ba^{-1}b = b^2a^{-1} = 0$. Moreover

$$e^{sI} = \sum_{n=0}^{\infty} \frac{s^n}{n!} I^n = 1 + sI$$

and

$$e^{sI}a = a + sb, \qquad e^{sI}b = b.$$

For any vector $x \in \Pi$, we define

$$\mathrm{rot}_s x = e^{sI}x,$$

which can be written

$$\mathrm{rot}_s x = e^{\frac{1}{2}sI}xe^{-\frac{1}{2}sI}. \tag{4.13.6}$$

The right hand side reduces to $x$ for any vector $x \perp \Pi$, since, as in the preceding cases, such a vector commutes with $I$. However, in contrast to the previous cases, an arbitrary vector in space, or more precisely, a vector $x \notin \Pi$ and not orthogonal to $\Pi$ cannot be decomposed into the sum of two vectors, one lying in $\Pi$, the other orthogonal to $\Pi$. The reason is that $\Pi$ is singular, $\mathrm{rad}\,\Pi = g \neq 0$. Nevertheless, even in the case of an arbitrary vector $x$ in space, formula (4.13.6) gives a correct expression for a parabolic rotation related to the light-like two-plane $\Pi$. Indeed, in view of $I^2 = 0$,

$$e^{\frac{1}{2}sI}xe^{-\frac{1}{2}sI} = (1 + \tfrac{1}{2}sI)x(1-\tfrac{1}{2}sI) = x + \frac{s}{2}(Ix - xI) - \frac{s^2}{4}IxI.$$

Setting $\frac{1}{2}(Ix - xI) = Tx$, $T$ is skew-symmetric (cf.(4.11.12)). Furthermore,

$$Tx = \tfrac{1}{2}(ba^{-1}x - xba^{-1}) = \tfrac{1}{2}\left\{ b(a^{-1}x + xa^{-1}) - (bx + xb)a^{-1} \right\}$$

$$= \frac{1}{(a \mid a)} \left\{ (a \mid x)b - (b \mid x)a \right\}.$$

This shows that $Tx \in \Pi$, that is $TE \subset \Pi$. In fact, $TE = \Pi$, because $(a \mid x)$ and $(b \mid x)$ simultaneously can be assigned arbitrary values. In particular, $Ta = b$, $Tb = 0$.

Remembering that $I^2 = 0$, we find

$$T^2x = \tfrac{1}{4}\left\{ I(Ix - xI) - (Ix - xI)I \right\} = -\tfrac{1}{2}IxI.$$

Hence

$$(4.13.7) \qquad e^{\frac{1}{2}sI}xe^{-\frac{1}{2}sI} = x + s(Tx) + \tfrac{1}{2}s^2(T^2x),$$

in complete agreement with (4.7.1).

A *common feature* of all the above rotations related to two-planes $\Pi$ of different kinds and expressed by formulas of the type

$$(4.13.8) \qquad x \longrightarrow e^{\frac{1}{2}\alpha I}xe^{-\frac{1}{2}\alpha I}$$

is that *they keep all vectors orthogonal to* $\Pi$ *unchanged*. This circumstance plays a basic role in the *superposition* of such rotations.

Let $E$ again be a vector space of Euclidean or of Lorentz signature, and $L$ an isometry in $E$, connected with the identity. Decompose $E$, according to 3.5 or 3.7, respectively, into mutually orthogonal subspaces $E_A$, $E_B$,... corresponding to certain specified factors of the minimal polynomial belonging to $L$. If such a subspace is of type $P$, that is a time-like plane, treated in 4.3 or in the present section, it is already a two-plane and nothing more has to be done about it. The same is true if the subspace is of the type $E_Q$, as treated in 4.2, and happens to be a two-plane. If dime $E_Q > 2$, $E_Q$ has to be decomposed into mutually orthogonal invariant two-planes, as is done towards the end of 4.8. A space $E_{-1}$, if present, is treated analogously (cf. 4.2). If a space $E_{+1}$, is present, it has to be treated differently, according as the corresponding factor of the minimal polynomial is $(t - 1)$ or $(t - 1)^3$. In the first case, we can neglect $E_{+1}$ entirely in the construction we are carrying out. In the second, we distinguish in $E_{+1}$ the invariant light-like two-plane $\Pi$, defined by the equation $(L - 1) E_{+1} = \Pi$.

The rotations related to all the invariant two-planes just mentioned can be expressed by formulae of the type (4.13.8) by means of the corresponding simple bivectors $I_M$. Hence, remembering what was said at the start of this superposition process, the isometry $L$ relative to the entire space $E$ can be expressed by

$$Lx = \prod_M e^{\frac{1}{2}\alpha_M I_M} x \left( \prod_M e^{\frac{1}{2}\alpha_M I_M} \right)^{-1} .$$

Finally, because the bivectors $I_M$, which belong to mutually orthogonal two-planes, all commute with each other, we obtain again

$$(4.13.9) \qquad Lx = e^F x e^{-F}, \text{ where } F = \tfrac{1}{2} \sum \alpha_M I_M .$$

4.14. The construction of the bivector $F$ figuring in the isometry formula (4.13.9) leads us to the following question. *Can any bivector* $F$ *be decomposed into the direct sum of mutually orthogonal simple bivectors?* As we can suspect, *the answer is affirmative in the cases of Euclidean or Lorentz signature and is negative for any other signature*.

Since bivectors, considered as operators by formula (4.11.10) are isomorphic to skew-symmetric transformations, we shall reformulate and treat the problem in terms of the latter.

Let $E$ be again a space regular with respect to a certain metric, and let $T_1$, $T_2$, $T$ be skew-symmetric transformations of $E$ into itself. We call $T_1$ and $T_2$ *orthogonal*, if the subspaces

$T_1E$ and $T_2E$ are orthogonal. A necessary and sufficient condition is that $T_1T_2 = 0$ and/or $T_2T_1 = 0$. To prove this, write $(T_1x \mid T_2y) = -(x \mid T_1T_2y) = -(T_2T_1x \mid y)$. The necessity of the condition follows from the supposed regularity of E, while the converse (in any of the forms stated) is obvious. We call T simple if TE is a two-plane. We will prove that in the Euclidean and in the Lorentz case every skew transformation S can be decomposed into the sum of mutually orthogonal simple skew transformations; and that this is not generally true for any other signature.

The proof of the equivalence of our assertions relative to bivectors and to skew transformations is left to the reader.

The affirmative part of our assertion concerning the decomposition of skew transformations can be based on our statement relative to the minimal polynomial of such transformations given at the end of section 4.8 and some other parts of the same section. We omit some quite easy arguments necessary to complete the proof.

We illustrate the decompositions in question by means of suitable bases. Let dim $E = n > 2$, with $n = 2\nu$ or $n = 2\nu + 1$ respectively. Denote by $e_1 e_2, \ldots, e_n$ any basis of E formed by mutually orthogonal unit vectors, with all $(e_k \mid e_k) = 1$ in the Euclidean case, and $(e_1 \mid e_1) = 1$, $(e_k \mid e_k) = -1$, for $k \neq 1$, in the Lorentz case. With a suitable basis, every bivector F, considered as an element of the corresponding Clifford algebra can be written on one of the following forms

$$(4.14.1) \qquad F = \sum_{k=1}^{\nu} \alpha_{2k-1,\, 2k}\, e_{2k-1}\, e_{2k}$$

or

$$(4.14.2) \qquad F = \beta\, (e_1 + e_2) e_3 + \sum_{k=2}^{\nu'} \beta_{2k,\, 2k+1}\, e_{2k} e_{2k+1} ,$$

with $\nu' = \nu$ if $n = 2\nu + 1$, $\nu' = \nu - 1$ if $n = 2\nu$ . (No $\Sigma$ occurs in (4.14.2) in the cases $n = 3$, $n = 4$.)

Only formula (4.14.1) can occur in the Euclidean case, while both can occur in the Lorentz case, excluding each other for any given bivector $F$. Still in the Lorentz case, in (4.14.1) the bivector $e_1 e_2$ is time-like, while the other bivectors are space-like. They are also space-like in the sum figuring in (4.14.2), while the bivector which figures in front of the sum is light-like. Indeed the vector $e_1 + e_2$ is light-like, the vector $e_3$ is space-like and orthogonal to $e_1 + e_2$.

Now we turn to the negative part of our statement and show first that the skew-symmetric transformation $T$ represented by the matrix (4.9.3) cannot be decomposed into simple skew-symmetric transformations orthogonal to each other. In section 4.9 the transformation $T$ acted on the coordinates, here we consider it as a transformation of a given basis.

Let $E$ be a four-dimensional vector space having the basis vectors $l_1, l_2, l_3, l_4$, To define a metric, we take the vectors $l_1, l_2, l_3, l_4$ isotropic, $(l_i \mid l_i) = 0$, set $(l_1 \mid l_4) = (l_2 \mid l_3) = 1$ and all the other scalar products $= 0$. Denote by $x = x^k l_k$ and $y = y^k l_k$ generic vectors in $E$, then we have

$$(x \mid y) = x^1y^4 + x^2y^3 + x^3y^2 + x^4y^1.$$

The space $E$ is regular and has signature $++--$, cf. 4.9.

Set

$$Tl_1 = l_1 + l_2, \quad Tl_2 = l_2, \quad Tl_3 = -l_3 - l_4, \quad Tl_4 = -l_4.$$

We easily find that $(l_i \mid Tl_k) + (l_k \mid Tl_i) = 0$, hence $T$ is skew-symmetric. The vectors $l_2$ and $l_4$ are eigenvectors with the respective minimal polynomials $(t - 1)$ and $(t + 1)$. The respective minimal polynomials of $l_1$ and $l_3$ are $(t - 1)^2$ and $(t + 1)^2$, and the minimal polynomial of the space $E$ is $m(t) = (t - 1)^2 (t + 1)^2$. The two-planes $(l_1, l_2)$, $(l_3, l_4)$, $(l_2, l_4)$ are invariant with the respective minimal polynomials $(t - 1)^2$, $(t + 1)^2$, $t^2 - 1$. Any other two-plane has a minimal polynomial of degree $\geqslant 3$, hence there are no other invariant two-planes (cf. the beginning of section 3.4).

We also note that the three planes just mentioned are totally isotropic and each of them is its own orthogonal complement.

If a skew-symmetric transformation $T$ can be decomposed, as above, into an orthogonal sum $\sum T_r$, then every subspace $T_kE$ is invariant under $T$. Indeed, because $T_kT_l = 0$ for $k \neq l$, we have $TT_k = T_kT_k = T_kT$, hence $T(T_kE) = T_kTE \subset T_kE$. Thus, in the case we are treating, the only possible $T_kE$ would be the above mentioned invariant planes $(l_1, l_2)$, $(l_3, l_4)$, $(l_2, l_4)$. No two of these planes being orthogonal to each other, the decomposition is proved to be impossible

For dimensions $> 4$ an analogous counterexample can be constructed by means of the following mild generalization of the preceding.

Let $E$ be a real vector space, regular with respect to a certain metric. Suppose there exists a skew-symmetric transformation $TE \subset E$, with a minimal polynomial of the form

$$m(t) = (t - \alpha)^2 (t - \beta)^2 t^p \prod_j (t^2 - \gamma_j) \prod_k (t^2 + \delta_k)$$

$\alpha \neq 0$, $\beta \neq 0$, $\alpha \neq \beta$, $p = 0, 1$, both the $\gamma_j$ and $\delta_k$ distinct positive numbers, $\gamma_j \neq \alpha^2$, $\gamma_j \neq \beta^2$. Such a transformation $T$ cannot be decomposed into the sum of mutually orthogonal simple skew-symmetric transformations.

The proof can be carried out on the same lines as in the four-dimensional case, but is slightly more complicated.

4.15. After this excursion into spaces with other signatures, we once more return to the Euclidean or Lorentz cases and, first of all, fill a gap left at the end of section 4.12. We are now able to show the following. Given a bivector $F$, one can, in general, find other bivectors $G$ such that $e^G = -e^F$; hence $e^{-G} = -e^{-F}$ and $e^G x e^{-G} = e^F x e^{-F}$. The only exceptions concern the Lorentz signature and are the following: the space of dimension 2, the case of a hyperbolic rotation in a space of dimension 3 and that of a parabolic rotation in spaces of dimensions 3 or 4.

In the Euclidean case such a $G$ can be constructed by adding $(2r + 1)\pi$ to one of the coefficients $\alpha_{2k-1,\, 2k}$

figuring in formula (4.14.1) or if $F = 0$ and the sum is empty, replacing $F = 0$ by $G = (2r + 1)\pi I$, with $I^2 = -1$. Analogous changes can be performed in the Lorentz case on any space-like bivector figuring in the sum in (4.14.1) or on any bivector figuring in the sum in (4.14.2). If the number of terms in the sums is less than $\nu$ or $\nu'$ respectively, a bivector $(2r + 1)\pi I$ can be added, such that $I^2 = -1$ and $I$ is orthogonal to the bivectors already occurring in the expression. The discussion of the exceptional cases is left to the reader.

Another application of the decomposition formulae (4.14.1) - (4.14.2) or of formula (4.13.9) is the following.

The exponential $e^F$ in (4.13.9) can be written in the form

$$(4.15.1) \qquad e^F = b_1 b_2 \dots b_{2h-1} b_{2h},$$

where $2h \leqslant n$, the vectors $b_k$ are unit vectors, i.e. $(b_k \mid b_k) = 1$ in the Euclidean case and $(b_k \mid b_k) = \pm 1$ in the Lorentz case, moreover, the two-planes spanned by the couples $b_{2p-1}$, $b_{2p}$ are mutually orthogonal. The products $b_{2p-1} b_{2p}$ commute with each other and are invariant under an arbitrary rotation performed in the corresponding two-plane.

Indeed, we have with the notations of section 4.13, for an elliptic rotation in either the Euclidean or in the Lorentz case

$$e^{\frac{1}{2}\theta I} = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\, ba^{-1}$$

$$= (\cos\frac{\theta}{2} . a + \sin\frac{\theta}{2}\, b)a^{-1} = b'b'',$$

with $(b' \mid b') = (b'' \mid b'') = \pm 1$.

We omit the analogous computation for a hyperbolic rotation and pass to the parabolic case. Here we have

$$e^{\frac{1}{2}sI} = 1 + \tfrac{1}{2}sba^{-1} = (a + \tfrac{1}{2}sb)a^{-1} = b'b''.$$

Since the light-like vector $b$ cannot be normalized, we may suppose without loss of generality that $(a \mid a) = -1$. Then we obtain $(b' \mid b') = (b'' \mid b'') = -1$.

Substituting (4.15.1) and the corresponding expression for $e^{-F}$ into (4.13.9), we again obtain formula (2.3.3), originally deduced by means of the Cartan-Dieudonné theorem. However, the formula now appears in a refined form, in view of the particular properties of the vectors $b_k$.

4.16. In the present section we make various additional remarks, mostly concerned with the bivectors of Clifford algebras.

Let $E$ be our ordinary vector space with an arbitrary metric and $C$ the associated Clifford algebra. An isometry $x \longrightarrow Lx$ induces an automorphism in $C$, whose effect on the product of a finite number of elements $\{x_k\}$ is given by the formula

$$L(\prod x_k) = \prod (Lx_k).$$

If the isometry can be expressed by an inner automorphism like (4.13.9), we have for this product and, as a matter of fact, for an arbitrary element $X$ of $C$, the simpler formula

$$LX = e^{F}Xe^{-F}.$$

We shall now show that the situation is analogous for infinitesimal isometries.

Let $T$ be a skew-symmetric transformation and thus $1 + \varepsilon T$

an infinitesimal isometry of the vector space $E$. Let, for example, $x, y, z$ be elements of $E$, and let us compute the effect of $T$ on the product $xyz$. We find in first approximation.

$$xyz \rightarrow (x+\varepsilon Tx)(y+\varepsilon Ty)(z+\varepsilon Tz)$$

$$= xyz + \varepsilon (Tx.yz + x.Ty.z + xy.Tz).$$

Hence

$$T(xyz) = Tx.yz + x.Ty.z + xy.Tz.$$

This formula becomes much simpler, if we express $Tx$ by a bivector as in 4.11. From

$$(4.16.1) \qquad Tx = Fx - xF$$

it will follow in fact that $T(\Pi x_k) = F.\Pi x_k - \Pi x_k.F$ and, for an arbitrary element $X$ of the Clifford algebra,

$$(4.16.2) \qquad TX = FX - XF.$$

It is sufficient to prove that $X \rightarrow X + \varepsilon (FX-XF)$ and $Y \rightarrow Y + \varepsilon (FY-YF)$ imply (in first approximation) that $XY \rightarrow XY + \varepsilon (FXY - XYF)$. This follows from

$$\{X + \varepsilon (FX - XF)\}\{Y + \varepsilon (FY - YF)\}$$

$$= XY + \varepsilon(FXY - XYF - XFY + XFY) = XY + \varepsilon(FXY - XYF).$$

The right hand side of (4.16.2) expresses the effect of an infinitesimal isometry on an arbitrary element of the Clifford algebra by an infinitesimal inner automorphism. Prove by reversion that a p-vector transforms into a p-vector!

A Lie algebra is a vector space of finite dimension over a field, for whose elements A, B, C,... there is given a law of multiplication $A \circ B$, with the usual distributive properties and satysfying the conditions $A \circ A = 0$ and $(A \circ B) \circ C + (B \circ C) \circ A + (C \circ A) \circ B = 0$. The last relation replaces the associative law.

It is easily seen that, if $A, B, C, \ldots$ are elements of an associative algebra, and $A \circ B$ is defined as $A \circ B = [A, B] = AB - BA$, then the above relations are automatically satisfied; the second reduces to the well known Jacobi identity

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0.$$

We show that *the bivectors of a Clifford algebra form a Lie algebra* under the product rule $F \circ G = [F, G]$. The only thing to be proved is that if $F$ and $G$ are bivectors then $[F, G] = FG - GF = H$ is also a bivector. A priori, $H$ could be the sum of a scalar, a bivector and a quadrivector, $H = S + B + Q$. By reversion (cf. p. 13) we obtain $\tilde{H} = \tilde{S} + \tilde{B} + \tilde{Q} = S - B + Q$. On the other hand $\tilde{H} = \tilde{G}\tilde{F} - \tilde{F}\tilde{G}$ $= (-G)(-F) - (-F)(-G) = -H$. This proves $S = 0$, $Q = 0$ and that $H$ is a bivector.

According to a general theorem of Lie, skew-symmetric transformations also form a Lie algebra, since they form the infinitesimal group of the group of isometries. Here is a direct proof. Let $T_1$ and $T_2$ be skew-symmetric, set $T = [T_1, T_2] = T_1 T_2 - T_2 T_1$; then $T$ is skew-symmetric. Indeed, $(x \mid Tx)$ $= (x \mid (T_1 T_2 - T_2 T_1)x) = (x \mid T_1 T_2 x) - (-1)^2 (T_1 T_2 x \mid x) = 0$.

We now write the correspondence $Tx = Fx - xF$ between

T and F in the form

$$T_F x = Fx - xF. \tag{4.16.3}$$

Then we have, in notations easy to understand,

$$[T_F, T_G] = T_{[F, G]}. \tag{4.16.4}$$

The straightforward proof is left to the reader. This identity expresses the fact that the spaces of skew-symmetric transformations and of bivectors are isomorphic also when considered as Lie algebras.

A Lie algebra including that of the bivectors is formed by the bivectors and vectors of a Clifford algebra. This is obvious by now. Let $x$ and $y$ be vectors, $F$ and $G$ bivectors. Then we have the following scheme: $[F,G]$=bivector, $[F,x]$=vector, $[x,y]$ =bivector.

4.17. We give here a short digression about some features of the theory of electromagnetic fields, considered in four-dimensional Lorentz space, the space-time of special relativity. With a slight change in our notations we use the metric form

$$\begin{gathered} (x|x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = g_{ik}x^i x^k \\ g_{00} = 1, \quad g_{11} = g_{22} = g_{33} = -1, \quad g_{ik} = 0, \; i \neq k. \end{gathered} \tag{4.17.1}$$

The coordinate $x^0$ corresponds to time, while $x^1$, $x^2$, $x^3$ are the ordinary coordinates used in three-space. As is well known, Minkowski has made a synthesis of the electric and magnetic vectors $E$ and $H$, uniting them in a skew-symmetric tensor or a bivector $F$. We denote the covariant and the contravariant components of $F$ by $F_{ik}$ and $F^{ik}$ respectively, the passage from the ones to the others being performed according to the ordinary rules of tensor calculus by means of the metric form (4.17.1).

For the passage from E and H to F we use the following dictionary

$$(4.17.2)\quad \begin{cases} \{E_x, E_y, E_z\} = \{F_{01}, F_{02}, F_{03}\}, \\ \{-H_x, -H_y, -H_z\} = \{F_{23}, F_{31}, F_{12}\}, \\ \qquad\qquad F_{ik} = -F_{ki}, \end{cases}$$

which implies that the diagonal terms vanish.

Set $\frac{\partial}{\partial x_k} = \partial_k$, then Maxwell's equations in their form due to Minkowski, read

$$(4.17.3)\quad \begin{cases} \partial_i F_{kl} + \partial_k F_{li} + \partial_l F_{ik} = 0, \\ \partial_k F^{ik} = -j^i, \end{cases}$$

where each now represents four equations. The $j^i$ are the contravariant components of the four-current j, which is considered as given and which must satisfy the condition $\operatorname{div} j = \partial_k j^k = 0$.

The first system of equations is usually integrated by means of a vector potential A related to F by the equations

$$(4.17.4)\qquad \operatorname{rot}_{ik} A = \partial_i A_k - \partial_k A_i = F_{ik}.$$

Usually A is also subjected to a supplementary condition, often called Lorentz' condition, namely

$$(4.17.5)\qquad \operatorname{div} A = \partial_k A^k = 0.$$

The equations (4.17.4) determine A only up to an additive term grad u. If moreover the supplementary condition (4.17.5) is to be satisfied, then u must be a solution of the homogeneous wave equation.

By means of (4.17.4) alone, the first of the systems

(4.17.3) is automatically satisfied (as alluded to before); by means of (4.17.4) and (4.17.5) the second system reduces to

(4.17.6) $$\Box A^i = j^i ,$$

where $\Box$ is the wave operator or the d' Alembertian

(4.17.7) $$\Box = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 .$$

In Clifford language Maxwell's equations take a remarkably simple form. If we introduce a basis $\{e_i\}$, that satisfies $(e_i | e_k) = g_{ik}$, and the dual basis $\{e^i\}$ satisfying $(e^i | e_k) = \delta^i_k$ (cf.(3.3.8)) and $(e^i | e^k) = g^{ik}$ ($= g_{ik}$ in this special case), we can consider the bivectors F as elements of the associated Clifford algebra and write

(4.17.8) $$F = \tfrac{1}{2}F^{ik}e_i e_k = \tfrac{1}{2}F_{ik}e^i e^k .$$

By means of the invariant differential operator nabla

(4.17.9) $$\nabla = e^k \partial_k$$

the equations (4.17.3) can be put into the condensed form

(4.17.10) $$\nabla F = j .$$

We only indicate the proof. Formula (4.11.9) for xF, with the components $x_k$ replaced by the operators $\partial_k$, gives $\nabla F$ as the sum of a vector and a trivector. Equation (4.17.10) expresses that the vector is equal to the given current vector j and that the trivector vanishes. The identity of the equations (4.17.3) and (4.17.10) follows from the quoted formula. The necessity of the condition div j = 0 follows from $\Box F = \nabla^2 F = \nabla j = \text{div } j + \text{rot} j$ (cf.(1.11.2)), and the fact that $\Box F$ and rot j are bivectors, while div j is a scalar.

The vector potential may be required to satisfy the equation

$$(4.17.11) \qquad \nabla A = F.$$

According to (1.11.2) this equation says div $A$ + rot $A = F$, that means div $A = 0$ and rot $A = F$. Hence (4.17.11) expreses both equation (4.17.4) and the supplementary condition (4.17.5). In view of $\nabla^2 = \square$, the equations (4.17.11) and (4.17.10) give $\square A = j$, that is the equations (4.17.6).

From Maxwell's equations we now turn to the so-called ponderomotive force or Lorentz force by which an electromagnetic field acts on a charged particle. Let such a particle, with charge $e$ and restmass $m$ describe a world line $x = x(s)$, that is a line with a time-like tangent. The parameter $s$ is supposed to be the arc length, which implies that the tangent vector $u = dx/ds$ satisfies the equation

$$(4.17.12) \qquad (u \mid u) = 1.$$

Hence the vector $u$ is the unit tangent vector to the world line. Setting the velocity of light $c = 1$, the equations of motion read

$$(4.17.13) \qquad m\frac{du_i}{ds} = f_i \quad \text{with} \quad f_i = eF_{ik}u^k .$$

Here the $du_i/ds$ are clearly the covariant components of the acceleration. The vector $f$ with the covariant components $f_i$ is called the Lorentz force. By (4.11.7) the last equation can also be written in the form

$$(4.17.14) \qquad m\frac{du}{ds} = -\tfrac{1}{2}\, e\, [\, u, F\, ] .$$

We will study these equations from the geometrical point of view. Recalling that $F_{ik}$ is skew-symmetric, the electromagnetic field, considered as an operator, defines in every point $x$ an infinitesimal isometry

$$(4.17.15) \qquad Tv_i = \frac{e}{m} F_{ik}(x)v^k$$

of the local vector space $\{v\}$ attached to the point $x$ (cf. (4.11.7)). For a moving point $x$, the tangent vector varies according to this isometry, in complete agreement with the fact that its length is constant ($= 1$).

The type of the infinitesimal isometry (4.17.15) depends on the local character of the bivector $F(x)$. In general, it can be decomposed into two infinitesimal rotations, one hyperbolic, the other elliptic, related to two mutually orthogonal two-planes, the one time-like the other space-like (cf. the decomposition formula (4.14.1)). However, there exist important classes of fields which induce parabolic rotations. Clearly, a necessary and sufficient condition is that the bivector $F(x)$ be light-like or, in other words, singular. The fields in question are called singular fields or null fields. We confine ourselves to showing that electromagnetic plane waves form a class of singular fields.

We mean here by a plane wave a solution $F(x)$ of the homogeneous Maxwell equations which is constant on a family of parallel three-planes. The equations of such planes are $(l \mid x) = \text{const.}$, where $l$ is a fixed vector and $x$ the position vector in space-time. We restrict ourselves to the most interesting case, namely $F_{jk}(x) = f_{jk} \cdot g((l \mid x))$, the $f_{jk}$ constant and $g$ a scalar function, independent of the subscripts. The vector potential $A(x)$ will be of the form $A(x) = a . h((l \mid x))$, $a$ being a constant vector and $h$ being a scalar function. This potential has to satisfy the equations

$\Box A(x) = 0$ and $\operatorname{div} A(x) = 0$, which give in the present case

$$(l|l).\ h''((l|x)) = 0, \quad (l|a).\ h'((l|x)) = 0.$$

The first equation requires either $h'' \equiv 0$ or $(l|l) = 0$. However $h'' \equiv 0$ would give a constant field, entirely uninteresting in the present context, where we are looking for a "wave". Discarding this case, we find the two conditions which are necessary and sufficient, namely

$$(4.17.6) \qquad (l|l) = 0, \quad (l|a) = 0.$$

The first equation tells us that the "phase vector" $l$ has to be light-like, and the second that the "amplitude vector" a has to be orthogonal to $l$, hence space-like. (These denominations are used here in a slightly generalized sense; they are usual in the exponential case treated below.) With $h' = g$ the bivector F can be written

$$(4.17.7) \qquad F = \tfrac{1}{2} g((l|x)).\ [l, a],$$

where the formula can be read either as a tensor formula or as a formula in Clifford language. With the latter meaning, we also can write the simpler formula

$$(4.17.8) \qquad F = g((l|x)).\ l a.$$

The covariant components of F are

$$(4.17.9) \qquad F_{ik} = g((l|x)).\ (l_i a_k - l_k a_i).$$

A monochromatic plane wave is obtained when $h((l|x)) = \exp i((l|x) + \delta)$, and the corresponding field becomes

$$F = ie^{i(l|x) + i\delta} l a = \tfrac{1}{2} i\, e^{i(l|x) + i\delta} [l, a].$$

If real quantities are wanted, one may take the real parts of the above expressions.

The essential fact for us is that the *bivector* F figuring in (4.17.7) or (4.17.8) *is singular*. In Clifford language this is expressed by the single equation $F^2 = 0$.

The singular character of F has the following well known implications. The electric and the magnetic vectors E and H have equal lengths and are orthogonal to each other and to the direction of propagation given by the three-vector $(l_1, l_2, l_3)$.

To verify this, we first remark that the substitution $a \rightarrow a - \lambda l$ neither changes the bivector $[l, a]$ nor the scalar product $(l \mid a)$. Hence we can suppose, without loss of generality, that the time component $a_o$ of a vanishes, that means that a is a vector in the physical three-space. In this case, according to our dictionary, we have except for a common scalar factor,

$$E_x = l_0 a_1 , \quad E_y = l_0 a_2 , \quad E_z = l_0 a_3 ,$$

$$- H_x = l_2 a_3 - a_2 l_3 , \quad - H_y = l_3 a_1 - a_3 l_1 \quad - H_z = l_1 a_2 - a_1 l_2$$

with

$$l_1^2 + l_2^2 + l_3^2 = l_0^2 \quad , l_1 a_1 + l_2 a_2 + l_3 a_3 = 0.$$

These equations imply our above assertions.

Returning to space-time and to the equations (4.17.13), we conclude that an electromagnetic plane wave moves a point charge in such a way that *the unit tangent vector to the world line undergoes an infinitesimal parabolic rotation* in every point of this line.

4.18. We now fill a gap left in 4.1 and explain in what sense the convergence of the series defining $e^A$ is to be understood.

Let $E$ be an n-dimensional real vector space. A set $M \subset E$ is called convex if $x \in M$, $y \in M$ implies $\alpha x + \beta y \in M$ for all $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$. It can be said equivalently that $M$ is convex if an only if

$$\lambda M + \mu M \subset (\lambda + \mu) M \tag{4.18.1}$$

for all positive $\lambda$ and $\mu$, where the left hand side is to be read in the sense of vectorial addition of the sets $\lambda M$ and $\mu M$. To prove this, let $x \in \lambda M$, $y \in \mu M$ and write

$$x + y = (\lambda + \mu) \left( \frac{\lambda}{\lambda + \mu} \frac{x}{\lambda} + \frac{\mu}{\lambda + \mu} \frac{y}{\mu} \right).$$

The points $\lambda^{-1} x$ and $\mu^{-1} y$ are in $M$, hence their weighted sum in the parenthesis is also in $M$, hence $x + y$ is in $(\lambda + \mu) M$.

A set $M \subset E$ is called symmetric if $x \in M$ implies $-x \in M$, or in other words if $M = -M$. In complex vector spaces the role of the symmetric sets is played by the circled sets. A set $M$ in a complex vector space is called circled if $x \in M$ and $|\lambda| = 1$ imply $\lambda x \in M$, or, in other words, if $\lambda M \subset M$ for $|\lambda| = 1$.

A set $M \subset E$ is called absorbing if for every $x \in E$ there exists a $\lambda > 0$ such that $\lambda x \in M$. Finally we call $M \subset E$ expelling if for every $x \neq 0$ in $E$ there exists a $\lambda > 0$ such that $\lambda x \notin M$.

Let now $M$ be a fixed convex, symmetric, absorbing and expelling set in $E$. For any vector $x \in E$ define $\| x \|$ as the greatest lower bound of all positive numbers $\lambda$ for which $x \in \lambda M$. This number $\| x \|$ is finite, since $M$ is absorbing.

Furthermore $\|x\|$ possesses the usual properties of a norm:

1) $\|x\| = 0$ if and only if $x = 0$;

2) $\|\lambda x\| = |\lambda| \cdot \|x\|$;

3) $\|x + y\| \leqslant \|x\| + \|y\|$.

Indeed 1) follows from the fact that $M$ is expelling, 2) from the convexity and the symmetry of $M$ and 3) from convexity using relation (4.18.1).

Let now $A$ be any endomorphism of $E$ and define $\|A\|$ as the greatest lower bound of all positive numbers $\lambda$ for which $AM \subset \lambda M$. $\|A\|$ is again finite, since $M$ is absorbing and $E$ finite dimensional. $\|A\|$ is a norm on the ring $\mathcal{E}(E)$ of all the endomorphisms of $E$, i.e. has the properties

1) $\|A\| = 0$ if and only if $A = 0$,

2) $\|\lambda A\| = |\lambda| \, \|A\|$,

3) $\|A_1 + A_2\| \leqslant \|A_1\| + \|A_2\|$,

4) $\|A_1 \; A_2\| \leqslant \|A_1\| \cdot \|A_2\|$,

5) $\|1\| = 1$.

Properties 1)-3) follow as above, while 4)-5) are obvious.

Furthermore we have for all $x \in E$ and all $A \in \mathcal{E}(E)$

(4.18.2) $$\|Ax\| \leqslant \|A\| \cdot \|x\| ,$$

This being an immediate consequence of the definitions.

The norms defined with the help of two different convex, symmetric, absorbing and expelling sets $M$ and $N$ are equivalent, i.e. there exist two positive numbers $\gamma > 0$, $\delta > 0$ such that

$\gamma \|A\|_M \leqslant \|A\|_N \leqslant \delta \|A\|_M$ for all $A \in \mathcal{E}(E)$.

follows from the obvious fact that there exist positive numbers $\lambda$ and $\mu$ such that $M \subset \lambda N$, $N \subset \mu M$.

From the completeness of the real and the complex fields and the finiteness of dim E follows that the space E and the space $\mathcal{E}(E)$ of the endomorphisms S are complete with respect to the norms just defined. This means that from $\| x_n - x_m \| \to 0$ follows the existence of an x such that $\| x - x_n \| \to 0$ and that from $\| S_n - S_m \| \to 0$ follows the existence of an S such that $\| S - S_n \| \to 0$.

Consider now a series $\sum_{i=0}^{\infty} A_i$ whose terms are elements of $\mathcal{E}(E)$. If the numerical series $\sum_{i=0}^{\infty} \| A_i \|$ is convergent, then $\sum A_i$ is said to be absolutely convergent and it follows from the completeness of $\mathcal{E}(E)$, just as in ordinary analysis, that there exists $S \in \mathcal{E}(E)$ such that $\| S - \sum_{i=0}^{n} A_i \| \to 0$ as $n \to \infty$. If this is the case, $\sum A_i$, is said to be convergent and S is called its sum. In particular we have then for every vector $x \in E$ by (4.18.2)

$$\sum_{i=0}^{\infty} A_i x = Sx.$$

Just as in ordinary analysis, the sum of an absolutely convergent series is independent of the order of the terms and two absolutely convergent series can be multiplied term by term. Since

$$\sum_{n=0}^{\infty} \frac{1}{n!} \| A^n \| \leq \sum_{n=0}^{\infty} \frac{1}{n!} \| A \|^n$$

is convergent, the series $\sum_{n=0}^{\infty} \frac{1}{n!} A^n$ with which we have defined $e^A$ in 4.1 is absolutely convergent.

These and similar considerations justify all the operations performed in 4.1 on the series $e^{sA}$, in particular the differentiations of arbitrary order with respect to the parameter s.

More generally, let E be a finite dimensional algebra over the real field. A set $M \subset E$ is called idempotent if $x \in M$, $y \in M$ implies $xy \in M$, or in other words if $MM \subset M$. If $M$ is any convex, symmetric, absorbing, expelling and idempotent set in $E$, we can for any $x \in E$ define $\| x \|$ as a greatest lower bound, as above. This norm will have the properties 1) - 3) and also the further property.

$$4) \quad \| xy \| \leqslant \| x \| \cdot \| y \| .$$

If, for example, for an element $c = \sum_A c_A e_A$ of a Clifford algebra we put $| c | = \sum_A | c_A |$, then the set of all elements with $| c | \leqslant 1$ forms a set $M$ of the above type.

Having defined this norm, we can consider absolutely convergent series in $E$ and in particular we can define the exponential $e^x$ for any $x \in E$. By this the introduction of $e^{sF}$ in section 4.11 is fully justified.

# BIBLIOGRAPHY

## I. Books

1. E. Artin, Geometric Algebra, Interscience Tracts in Pure and Applied Mathematics No. 3, New York 1957, Interscience Publishers.

2. E. Cartan, Nombres Complexes, Encyclop. Sc. math., French edition, I. 5, 1908 = Oeuvres complètes, Partie II, Paris 1953, pp. 107-246, Gauthier-Villars.

3. E. Cartan, Leçons sur la théorie des spineurs, I. Les spineurs de l'espace à trois dimensions, Actualités Scientifiques et Industrielles No. 643, Paris 1938, Hermann.

4. E. Cartan, Leçons sur la théorie des spineurs, II. Les spineurs de l'espace à $n > 3$ dimensions, les spineurs en géometrie riemannienne, Actualités Scientifiques et Industrielles No. 701, Paris 1938, Hermann.

5. C. Chevalley, Theory of Lie groups I, Princeton Mathematical Series No. 8, Princeton 1946, Princeton University Press.

6. C. Chevalley, The algebraic theory of spinors, New York 1954, Columbia University Press.

7. C. Chevalley, The construction and study of certain important algebras, Publications of the Mathematical Society of Japan, No. 1, Tokyo 1955.

8. J. Dieudonné, La géométrie des groupes classiques, Ergebnisse der Mathematik und ihrer Grenzgebiete, Neue Folge, Heft 5, Berlin-Göttingen-Heidelberg 1955, Springer-Verlag.

9. P. K. Raševskiĭ, The theory of spinors, Uspehi Mat. Nauk (N.S.) vol 10, No. 2 (64), pp. 3-110 (1955) = Amer. Math. Soc. Translations, Series 2, vol. 6, pp. 1-110 (1957).

## II. Papers

10. W. K. Clifford, Application of Grassmann's Extensive Algebra, Amer. Journ. of Math., vol. 1 (1878) pp. 350 - 358 = Mathematical Papers, London 1882, pp. 266-276.

11. R. Lipschitz, Comptes Rendus Acad. Sc. Paris, vol. 91 (1880), pp. 619, 660 = Bulletin des Sciences Math., (2) vol. 11 (1887), p. 115.

12. M. Riesz, Sur certaines notions fondamentales en théorie quantique relativiste, Dixième Congrès des Mathematiciens Scandinaves, Copenhague 1946, pp. 123-143.

13. M. Riesz, L'Equation de Dirac en Relativité Générale, Douzième Congrès des Mathematiciens Scandinaves, Lund 1953, pp. 241 - 259.

## III. Additional references on tensor algebras

14. N. Bourbaki, Eléments de Mathématique, I. Partie, Livre II: Algèbre, Chapitre III: Algèbre multilineaire. Paris 1948, Hermann.

14. A. Lichnerowicz, Eléments de calcul tensoriel, Collection Armand Colin No. 253, Paris 1951.

# Index

Absorbing 184
Anticommute 8
Associativity 9
Automorphism of a metric 73, 95
Basis 9, 34
Bivector 21, 67
Canonical decomposition 101
Cartan-Dieudonné theorem 2, 74
Center 14
Central idempotents 17
Chevalley 1, 38, 80, 81
Circled 184
Clifford algebra 8, 31, 38, 156
Clifford group 80
Clifford number 8
Collinear vectors 27
Commutator 28, 31
Complementary bases 94
Connected component of the identity 132
Convex 184
Decomposition of a vector space 86
Derivation 38, 39
Divided commutator 30, 31
Eigenvalue 84
Eigenvector 84
Elliptic rotation 113
Empty product 9
Endomorphism 84
Euclidean metric 7, 73, 91, 100
Expelling 184
Exterior algebra 7
Fueter 37
Gårding 107
Graded algebra 27, 38
Grassmann algebra 7
Hamilton 2, 71
Hyperbolic rotation 115
Ideal 14
Idempotents 17
Infinitesimal isometry 128
Infinitesimal transformation 131
Invariant subspace 84
Inverse 89
Invertible 77

Irreducible subspace 96
Isometry 73, 95
Isotropic 73, 90, 105
Kronecker symbol 6
Laplace operator 5, 35
Length 32
Lie algebra 175
Light cone 105
Light-like subspace 107
Light-like vector 105
Linear transformation 84
Lorentz force 180
Lorentz metric 7, 73, 91, 105
Lorentz transformation 115
Main antiautomorphism 13
Main involution 13, 39
Maxwell's equations 178
Metric 6, 8, 22, 90
Minimal polynomial 85
Minkowski 178
Monochromatic plane wave 182
Nabla 35, 179
Norm 185
Odd 13
Orthogonal 7, 27, 73, 90
Orthogonal basis 23
Orthogonal complement 92
Orthogonal projection 93
Orthogonal skew-symmetric transformation 168

Parabolic rotation 119,139
Plane wave 181
Ponderomotive force 180
Pseudoscalar 11, 13, 31
p-vector 30
Quaternion 20
Quotient space 92
Radical 92
Reciprocal polynomials 98
Reducible subspace 96
Reflection 79, 102
Regular metric 24, 73, 90
Reversion 13
Rotation 79, 102
Scalar 8, 13, 28
Scalar product 6, 8, 22
Semi-simple algebra 19
Simple algebra 15
Simple p-vector 30
Simple skew-symmetric transformation 169

Singular metric 24, 90

Skew 128

Skew-symmetric linear transformation 128

Skew-symmetric square root of (-1) 133

Skew-symmetric tensor 28

Space-like subspace 107

Space-like vector 105

Spectral decomposition 89

Sylvester's law of inertia 149

Symmetric 184

Symmetry 74

Tensor algebra 27

Time-like subspace 107

Time-like vector 105

Totally isotropic 91

Vector 8, 13, 21, 28

Vector potential 178

Wave operator 179

# Contents

Preface 1

I. Clifford Numbers 5

II. Rotations and Reflections 71

III. Canonical Representation of Isometries 84

IV. Representation of Isometries by Infinitesimal Transformations and Clifford Bivectors 128

Bibliography 188

Index 190

U of Maryland Students' Supply Store 72567

University of
Connecticut
Libraries